

TARTU ÜLIKOOL
Arvutiteaduse instituut
Informaatika mitteinformaatikutele

Mari Väljaots
Avaandmete avaldamise mõjude hindamise
töövahend

Magistritöö (15 EAP)

Juhendajad: Vambola Leping, MSc
Ott Velsberg, MSc

Tartu 2019

Avaandmete avaldamise mõjude hindamise töövahend

Lühikokkuvõte

Magistritöö eesmärk on pakkuda avaandmete avaldamise mõjude hindamise töövahend, mis toetaks teabevaldajat otsuse, kas andmeid tohib avaandmetena avaldada või mitte, tegemisel. Eesmärgi saavutamiseks kaardistas töö autor hetke olukorra avaandmete valdkonnas. Ülevaate saamiseks tehti lõppenud arendusprojektide analüüs ja viidi läbi intervjuud avaandmete spetsialistidega.

Avaandmete avaldamise mõjude hindamise töövahend sisaldab küsimusi, millele teabevaldaja peab enne andmete avaldamist mõtlema. Küsimustik aitab hinnata olulisi nüansse avaandmete kontekstis.

Loodud töövahend võiks olla sisendiks avaandmete valdkonna poliitika kujundajatele, kes saavad magistritöö raames analüüsitud info võtta sisendiks valdkonna edasisel arendamisel.

Võtmesõnad: avaandmed, avaliku sektori andmed, mõjude hindamine

CERCS: P170 Arvutiteadus, arvanalüüs, süsteemid, juhtimine (automaatjuhtimisteooria)

Impact Assessment Tool for Open Data Publication

Abstract

The aim of this Master's Thesis is to provide a tool for public sector information holders with the aim to support them evaluate the impact of opening data and decide whether or not to publish data as open data. To achieve this goal, the author mapped the current situation in the field of open data. An overview of completed development projects was conducted to get an overview of current status and interviews were conducted with open data specialists.

Open data publication impact assessment tool contains guiding questions that the holder of the data needs to pay attention to before publishing the data. The questionnaire helps to evaluate important nuances in the context of open data.

The created tool can be an input for policy makers in the field of open data who can take the information analyzed in the thesis as input for further development of the field.

Keywords: open data, open government data, impact assessment

CERCS: P170 Computer science, numerical analysis, systems, control

Sisukord

1	Sissejuhatus.....	4
2	Teooria ülevaade.....	6
2.1	Mis on avaandmed?.....	6
2.2	Avaandmete mõju	7
2.3	Metaandmete tähtsus	9
2.4	Juhised avaandmete avaldamiseks	9
2.5	Avaandmete mõjuhindang.....	9
2.6	Rahvusvaheline praktika avaandmete valdkonna toetamiseks.....	10
2.7	Arendusprojektide toetusmeetmed.....	11
3	Metoodika	13
3.1	Andmekogumise meetod.....	13
3.2	Andmete analüüsi meetod	14
4	Projektide ja intervjuude analüüs.....	16
4.1	Projektide analüüsi tulemus	16
4.2	Intervjuude analüüsi tulemus	17
4.2.1	Avaandmete avaldamise kohustus	17
4.2.2	Huvi avaandmete kasutamise osas.....	18
4.2.3	Avaandmete avaldajad	19
4.2.4	Tagasiside avaandmete portaalile	19
4.2.5	Avaandmete tervikvaade.....	19
4.2.6	Avaandmete litsents	20
4.2.7	Kokkuvõte intervjuudest.....	20
5	Avaandmete avaldamise mõjude hindamise töövahend	21
5.1	Ülevaade avaandmete avaldamise mõjude hindamise töövahendist	21
5.2	Ülevaade avaandmete avaldamise mõjude hindamise protsessist.....	22
5.3	Tagasiside avaandmete avaldamise mõjude hindamise töövahendile.....	26
6	Tulemused.....	27
7	Kokkuvõte.....	29
8	Viidatud kirjandus	30
9	Lisad	33
9.1	Lisa I Intervjuu kava	33
9.2	Lisa II Avaandmete avaldamise mõjude hindamise töövahend	35
9.3	Lisa III Avaandmete avaldamise mõjude hindamise protsessi terviklik skeem.....	38
10	Litsents.....	39

1 Sissejuhatus

Eesti on edukas ja kiire arenguga digiriik. Eesti tehnoloogilised lahendused on innovatiivsed ja pakuvad huvi ka väljapoole. Siia tullaakse üle maailma, et õppida Eesti eduloost, et seejärel naasta koduriiki e-teenuseid arendama ning e-valitsemist juurutama. Eesti on suunanäitajaks oma x-tee, ID-kaardi lahenduste, esimesena loodud andmesaatkonna, e-residentsuse programmiga ning palju muuga. Euroopa Komisjoni 2018. aasta digitaalmajanduse ja -ühiskonna indeksi (DESI) järgi on Eesti Euroopa Liidus 9ndal kohal [1]. DESI aruandes antakse ülevaade liikmesriikide edusammudest digiteerimisel, milles Eesti on alates taasiseseisvumisest alates olnud edukas. Ühendatud Rahvaste Organisatsiooni (ÜRO) e-valitsemise arengu indeksis on Eesti 2018. aastal 16. kohal 193 riigi konkurentsis [2]. 2018 aasta tulemus on kolm kohta halvem kui aastal 2016, kui saavutati 13. koht.

Eestis on asjaajamine riigiga peaaegu¹ täielikult võimalik elektroonselt. DESI indeksi järgi on Eesti aastaid olnud esimeste seas internetipõhiste avalike teenuste osutamisel ning lausa 96% elanikest kasutab e-teenuseid, mis on Euroopa Liidu keskmisest kaks korda kõrgem [1]. Samuti on riik üheks suuremaks andmete kogujaks läbi erinevate infosüsteemide ja domeenide. Sellest olenemata on üks valdkond digiriigis jäänud tagaplaanile. Olenemata avaliku teabe seaduse (AvTS) järgi tulenevast kohustusest, on avalik sektor andmete avaldamisel teinud tagasihoidlikku tööd. Üks käesoleva magistr töö eesmärgi on välja selgitada selle põhjus.

2014. aastal kinnitas ja avaldas vabariigi valitsus „Eesti avaliku teabe masinloetava avalikustamise roheline raamat“ (avaandmete roheline raamat). Avaandmete roheline raamat annab suunised ja juhised, kuidas ning mida on vajalik avaldada. Kuigi Eestis on alates aastast 2015 kohustus teha andmed masinloetavalt kättesaadavaks, pole seda nõuet kõik avaliku sektori asutused täitnud. DESI 2018. aasta indeksis on Eesti avaandmete osas võrreldes 2016. aastaga tõusnud ning olukord on paranenud, kuid kasvuruumi on, sest avatud andmete kättesaadavus koguanalüüsides on 58%, mis seab Eesti Euroopa Liidu riikide seas 23. kohale [1]. Maailma mastaabis oli Eesti avaandmete valdkonnas 2017. aasta seisuga 44 kohal [3].

Läbipaistvuse huvides on oluline, et avaldatakse kõikvõimalikke andmeid, millel puuduvad seadustest tulenevad piirangud ning nende edasisel kasutamisel puuduvad samuti piirangud. Avaandmete puhul on oluline, et need oleks avaldatud masinloetavas formaadis. Riigil on palju andmeid, mille analüüsimine annab uut teadmist ning võimalusi teha ümberkorraldusi, kui see peaks olema vajalik. Selleks, et neid andmeid oleks mugav koostoides analüüsida, peaksid nad olema avaldatud viisil, mis teeks andmete kasutamise masinloetaval kujul võimalikult mugavaks.

Avaandmete avaldamise eel on kohustuslik läbi viia avaandmete mõjuhindang [4]. Mõjuhindang on vaja teha andmekogude põhimääruse muutmisel ja uue andmekogu loomisel, kui on soov andmeid avaldada avaandmetena. See kohustus ei ole täna rakendunud ning vastavaid mõjuhindanguid pole suuresti teostatud.

Käesoleva magistr töö eesmärk on luua praktiline vahend avaandmete avaldamisele eelneva mõjuhindangu tegemiseks. Antud praktiline vahend aitab teabevaldajat otsuse tegemisel, kas andmekogus sisalduvaid andmeid võib avaldada või mitte ning kas enne avaldamist oleks vaja andmeid anonümiseerida.

Magistr töö eesmärkide saavutamiseks on sõnastatud kolm uurimisküsimust:

- 1) Miks on avaldatud avaandmekogude arv väike?
- 2) Kuidas edendada avaandmete avaldamist?

¹ Teadlikult on jäetud abiellumine, lahutamine ja kinnisvara ost-müük mitte e-teenusteks

3) Kas ja kuidas avaandmete mõjuhinna koostamise vahend toetab andmete avaldamist?

Magistritöö teine peatükk selgitab, mis on avaandmed ja nende väärtus, mis on valdkonna regulatsioonid ja milline on rahvusvaheline praktika avaandmete avaldamise edendamisel. Kolmas peatükk sisaldab töö uurimismetoodika kirjeldust, mida ja kuidas analüüsitakse. Neljas peatükk sisaldab projektide ja intervjuude analüüsi tulemusi. Viiendas peatükis antakse ülevaade avaandmete avaldamise mõjude hindamise töövahendist. Kuuendas peatükis on magistritöö järeldused ja tulemused. Viimane seitsmes peatükk võtab magistritöö kokku. Lisaks on ülevaade kasutatud materjalidest ning neli lisa, sh töö litsents.

Magistritöö autorit ajendas töö teemat valima huvi avaandmete valdkonna vastu, kuid ka võimalus panustada magistritöö raames tehtava uurimusega igapäevasel töökohal. Magistritöö autor töötab kirjutamise ajal Majandus- ja Kommunikatsiooniministeeriumis riigi infosüsteemide osakonnas, mille üheks vastutusalaks on avaandmete poliitika eestvedamine (ei kuulu autori tööülesannete hulka). Eesti seis eelnevalt nimetatud uuringute, indeksite põhjal avaandmete valdkonnas on pigem kehv, mistõttu on vaja teada saada selle põhjused ning pakkuda lahendus valdkonna arendamiseks.

2 Teooria ülevaade

2.1 Mis on avaandmed?

Avaandmete (*open data*) all mõistetakse kõigile vabalt ja avalikult kasutamiseks antud masinloetavas formaadis andmeid, millel puuduvad kasutamist ning levitamist takistavad piirangud [5]. See tähendab, et avaldatud andmetele on lisatud litsents, mis ei sea andmete kasutamiseks piiranguid ning igaüks võib neid kasutada ja soovi korral rakendustes kasutada.

AvTS² sätestab, et avalik sektor peab avaandmeid avaldama teabevärava kaudu, milleks on riigiportaal www.eesti.ee. Riigiportaal omakorda suunab info otsijad edasi Eesti riigi avaandmete portaali³. Avaandmete portaal on igaühel ligipääs avaliku sektori juurdepääsupiiranguteta andmetele koos nende taaskasutamise ja edasijagamise õigusega nii äri- kui ka mitteäri-eesmärkidel. Riigi avaandmete portaali on mõeldud Eesti avaliku sektori asutuste kogutavate andmete avalikustamise hõlbustamiseks ja nende leidmise kergendamiseks potentsiaalsete kasutajate jaoks. Samuti saavad portaali kaudu avaldada enda Eesti kohta käivaid avaandmeid ka kõik teised, kes on end kasutajaks registreerinud. Avaandmete portaal saab avaandmeid otsida ja neid alla laadida, lisada viiteid uutele andmetele, otsida ja kasutada avaandmete põhjal loodud rakendusi, algatada avaandmete teemal arutelusid ning postitada teateid, küsimusi ja juhendeid.

Näiteks on avaandmed avaliku korra vastased ja valikus kohas toimepandud süütegude andmed, mida avaldab Politsei- ja Piirivalveamet eesmärgiga muuta Eesti turvalisemaks. Samuti on avaandmetena Tartu linn avaldanud spordi- ja mänguväljakute asukohad.

Kõik avalikud andmed AvTS järgi ei ole automaatselt avaandmed. AvTS § 29 lõige 3 sätestab andmed, mis on Eestis kohustuslik avaandmetena avaldada. Kuigi on üldine andmete avaldamise kohustus, siis sama seadus täpsustab, et andmeid tuleb avaldada, kui see on võimalik ja asjakohane [4]. Avaandmetele tuleb teabevaldajatel juurdepääs võimaldada failivormingus, mis on struktureeritud selliselt, et tarkvararakendused suudavad spetsiifilisi andmeid, sealhulgas üksikuid faktiväiteid ja nende sisemist struktuuri kergesti tuvastada, ära tunda ja välja lugeda [6]. Andmed tuleb avaldada masinloetaval kujul ning platvormist sõltumatus vormingus [4]. Andmete taaskasutamine tuleb üldsusele kättesaadavaks teha avatud vormingus. Kui avaandmete digitaalsele kujule, masinloetavale kujule või avatud vormingusse viimine ei ole võimalik või nõuab ebaproportsionaalselt suuri pingutusi, siis peab teabevaldaja võimaldama avaandmete juurdepääsu nende algkujul või mis tahes muus vormis [4]. Avaandmeteks ei loeta juurdepääsupiiranguga isikuandmeid ja andmeid, mille levik on seadusega piiratud. AvTS järgi kehtestab konkreetse valdkonna eest vastutav minister täpse loetelu, mis isikuandmed avaandmetena avaldamisele ei kuulu.

Avaandmed on:

- tasuta kättesaadavad kõigile isikutele mistahes kasutuseesmärgil,
- digitaalsed, masinloetavad ja riskasutatavad teiste andmetega,
- litsentseeritud kitsendusteta kasutamiseks ja edasilevitamiseks [5].

2007 aasta 7-8 detsembril kogunesid 30 avatud valitsemise (*open government*) eestkõnelejat Californiasse ja leppisid kokku kaheksas põhimõtte, mis käsitlevad avaandmeid [7]. Eesti avaandmete poliitika lähtub samadest põhimõtetest ning need on kajastatud avaandmete rohelises raamatus. Kaheksa kriteeriumit, mille järgimisel ja täitmisel on tegemist avaandmetega [7]:

² Avaliku teabe seadus, <https://www.riigiteataja.ee/akt/115032019011?leiaKehtiv>

³ Eesti avaandmete portaali aadress: <https://opendata.riik.ee>

1. Terviklikkus. Kõik avalikud andmed, mida ei piira isikuandmete ja muul põhjusel salastatud andmete piirangud, tehakse kättesaadavaks.
2. Pärinemine algallikast. Andmed on töötluseta kogutud algallikast, säilitades oma originaalkuju ja detailsuse. Põhimõte tuleb andmekogudest, kus samuti pole lubatud andmete võtmine teisest andmekogust. Teiste andmekogude andmeid on lubatud lisada andmehulka, kuid neid ei tohi kolmas osapool mitmendat korda taaskasutada.
3. Ajakohasus. Andmehulk on avaldatud võimalikult kiirelt, et säilitada selle ajakohasus.
4. Kättesaadavus. Andmed on tasuta kättesaadavad võimalikult paljudele avaandmete kasutajatele võimalikult laia kasutuseesmärgiga.
5. Masinloetavus. Andmed on mõistetava struktuuriga ja automaatselt töödeldavad.
6. Diskrimineerimise vältimine. Andmed on esitatud avalikult, nende kättesaamiseks pole vaja registreeruda ega taotleda juurdepääsu.
7. Avatud standardite kasutamine. Andmed on esitatud avatud formaadis, mis ei ole ühegi ettevõtte ega isiku ainuomand.
8. Vaba litsents. Andmed ei ole kaitstud autoriõiguse, patendi, kaubamärgi ega ärisaladuse regulatsiooniga.

Lähtuvalt magistritöö eesmärgist, keskendutakse käesolevas töös avaliku sektori avaandmetele ning nende avaldamisele. See ei tähenda, et muud andmed ei ole või ei saaks olla avaandmed.

2.2 Avaandmete mõju

Janssem, Charalabidis ja Zuiderwijk hinnangul on avaandmed väärtuslikumad ja loovad rohkem kasu, kui seda võimaldaks andmekogumite müümine [8]. Austraalias on andmete avalikult kättesaadavaks tegemise kohta tehtud põhjalik kulu-tulu analüüs, mis tõestab, et avalikult andmete kättesaadavaks tegemine on riigile kasulikum kui andmete müümine [9]. Näitena analüüsist saab tuua ruumiandmete avaldamise, mille avaldamise kulud olid 13 korda väiksemad võrreldes sellega, mis riigile tuluna hiljem tagasi tuli [9].

Avaandmetest saavad kasu mitmed valdkonnad tervikuna, näiteks poliitika, majandus, sotsiaalvaldkond jne, aga ka erinevate valdkondade spetsialistid, nagu teadlased, ettevõtjad, ajakirjanikud. Suuresti on kasu avaandmetest ka ametnikele endile, kelle töö on andmeid avaldada. Masinloetavalt on võimalik analüüsida oluliselt suuremat andmemahutu ja teha järeldusi ja otsuseid võrreldes mitte masinloetava viisiga ehk erinevaid dokumente läbi töötades. Ühendkuningriigi tehnoloogiajuht Liam Maxwell suutis 2010. aastal avaandmeid kasutades leida riigieelarvele säästu 4 miljoni naela ulatuses [10]. Ühendkuningriigis on avaandmetena kättesaadavad kuluaruanded, mida analüüsides selgus, et erinevad avaliku sektori organisatsioonid tellivad samasisulisi uuringuid, raporteid ning kui dubleerimine lõpetada ja omavahel infot vahetada tellitud raportitena, saab riigieelarves 4 miljoni naela ulatuses panustada muudeks tegevusteks [10].

Analüüsides erinevaid avaandmeid, on võimalik luua uusi avalikke teenuseid, mis on vajalikud erinevatele ühiskonnagruppidele ning võimaldavad asjaajamist muuta kiiremaks ja mugavamaks. Näiteks on Tallinna Sadam oma avaandmetest loonud lahenduse, millest on võimalik reaalselt näha, kui palju ning mis riikidest on pärit inimesed, kes Eestit läbi sadama

külastavad⁴. Samuti aitavad avaandmed muuta riigivalitsemist läbipaistvamaks ja tõsta poliitiliste otsuste kvaliteeti läbi analüüsiprotsessi. Selleks, et inimeste elusid mõjutavaid poliitilisi otsuseid vastu võtta, ei saa nende eelduslike mõjude ja kasu osas vale infot anda, sest kõigil on ligipääs andmetele, mille alusel otsuseid põhjendatakse. Inimestel on võimalus samasuguseid analüüse huvi korral järgi teha. Kindlasti on avaandmed olulised ka ülikoolide jaoks, võimaldades teaduslikke uuringuid jms läbi viia.

Avaandmete üheks eesmärgiks on elavdada majandust. Avaandmete kasutamisel on võimalik luua innovaatilisi rakendusi, mis võivad igapäevast elu muuta lihtsamaks. Näiteks analüüsides koos nii transpordi, liikluse kui ka inimeste liikumise andmeid, saab ühistranspordi graafikud parandada lähtuvalt elanike vajadustest ja luua ning kasutusele võtta targa linna (*smart city*) lahendusi. Laursen ja Salter ütlevad, et innovatsiooniliste lahenduste loomisel on avaliku sektori avaandmed ettevõtjatele väärtuslikud, sest aitavad kokku hoida aega ja kulusid uuenduslike teenuste ja/või toodete arendamisel, loomisel [11]. Selleks, et avaandmete abil igapäevaelu lihtsustavaid rakendusi luua, peavad avaldatud andmed olema ajakohased. Oluline on regulaarselt avaandmete portaalis andmeid uuendada.

Erinevad autorid nimetavad suhteliselt sarnaseid asju avaandmete kasuna. Siiski on oluline vaadata iga andmehulga pealt, mida on võimalik nende andmetega või koosmõjus mõne teise andmehulgaga teha, luua ja saavutada. Seega on oluline alati hinnata kasu konkreetsete andmete pealt, mitte üldistades tervele nõ umbmäärasele avaandmete hulgale. Avaandmed iseseisvana ei ole kasulikud, neid kasutades selgub nende täielik väärtus ja kasu laiemale ühiskonnale. Oluline on ka arvesse võtta, et enamus avaandmetest, mida taaskasutatakse, ei ole isikuandmed.

Eesti Infoühiskonna arengukava 2020 ühe suure eesmärgina on plaanis avalike e-teenuste kvaliteedi ja kasutajakogemuse järgmise arenguhüppena arendada välja asutusteüleselt toimivaid sündmuspõhiseid teenuseid [12]. Inimene peab saama asjad aetud ühe suhtluskorraga, igal võimalusel pigem automaatselt ja sekkuma ainult riigi poolse algatusel või märguandel. Eesmärgiks on muuta avalikud teenused kodanikele võimalikult lihtsaks. Selleks, et Eesti Infoühiskonna arengukava 2020 eesmärki saavutada, on oluline roll äriprotsesside uuendamisel. Magalhaesa ja Roseirab on leidnud, et avaandmeid saab kasutada nii era- kui ka avalikus sektoris konkurentsieelise loomiseks ning näitena on nimetatud toodete ja teenuste loomist, aga ka äriprotsesside parandamist, uuendamist [13]. Selleks, et edukalt ja valdkondade üleselt minna sündmuspõhiste teenuste osutamisele, on tähtsal kohal õigesti tehtud äriprotsessilised muudatused ning siin saavad samuti olla abiks avaandmed. Oluline on osata olemasolevaid andmeid kasutada ning nende väärtust mõista protsesside uuendamisel ja innovatsiooni loomisel.

Kuigi rohkem tuuakse välja avaandmetest saadavat võimalikku kasu, on kirjeldatud ka takistusi avaandmete avaldamise teemal. Erinevad autorid on artiklites kirjeldanud avalikus sektoris nimetatud põhjuseid, miks avaandmeid ei avaldata. Janssen, Charalabidis ja Zuiderwijk [8]; Conradie ja Choenni [14]; Barry ja Bannister [15]; Young ja Yan [16] on kirjeldanud, et valitsuse tasandil on kõige enam põhjuseks toodud järgmist:

- puudub tehniline infrastruktuur
- puudub asjatundlikkus andmete valdkonnas, kes tegeleks avaandmete avaldamisega
- olemasolevate andmete kvaliteet ei võimalda nende avaldamist – ehk oleks vaja tegeleda esmalt andmete korrastamisega, kui tulles tagasi eelnevalt nimetatud probleemi juurde ei ole inimest, kes teemaga tegeleks
- ei ole kindlustunnet, et andmete avaldamisega turvariske ei kaasne või kardetakse eksida konfidentsiaalsusnõuete vastu

⁴ BNS „Tahad teada, millistest riikidest Eestisse tullakse? See kaart näitab sadama läbijaid reaajas“
<https://raha.geenius.ee/rubriik/uudis/tahad-teada-millistest-riikidest-estisse-tullakse-see-kaart-naitab-sadama-labijaid-reaajas/>

- teema ei ole prioriteetne, millele peaks ressursse kulutama ja tegelema.

2.3 Metaandmete tähtsus

Andmetel ei ole väärtust, kui pole teada, mis andmetega on tegu. Seetõttu on oluline, et andmete kohta on olemas andmed ehk metaandmed. Avaandmete kontekstis on oluline, et andmete metaandmed oleksid samuti avaandmetena avaldatud. Kvaliteetsed metaandmed on olulised süsteemide jaoks, mis kasutavad andmeid erinevatest allikatest või tegelevad andmekaevega [16]. Oluline on teada, kas tegemist on inimeste poolt käsitsi loodud andmetega või on andmed kogunud masinad. Teadmata andmete tausta ja tähendust, pole neid võimalik õigesti kasutada. Metaandmed on olulised nii inimestele kui ka masinatele. Kumbki osapool ei saa andmeid ilma metaandmeteta kasutada, sest pole teada, mis andmetega on tegu. Kuigi standardid ja andmete avaldamise vormid võivad erineda, siis metaandmed on olulised ning annavad andmete kasutajatele teadmise, mis andmetega on tegemist ja seetõttu on tähtis, et metaandmete standardeid järgitaks [16].

2.4 Juhised avaandmete avaldamiseks

Avaandmete avaldamiseks on koostatud kaks soovituslikku juhend-dokumenti, mis on teabevaldajatele abiks avaandmete avaldamisel. Riigi Infosüsteemi Ameti poolt on koostatud 2015 aastal juhend „Avaandmete loomise ja avaldamise juhend⁵“, mida on uuendatud 2016. aastal. Tegemist on põhjaliku ülevaatega, kuidas avaandmeid avaldada ning millele tähelepanu pöörata, sh ka peale avaandmete avaldamist. AvTS ütleb, et andmed tuleb avaldada masinloetavalt. Täiendavad soovitusel ja ettepanekud, kuidas avaandmeid avaldada, on kirjas avaandmete rohelises raamatus. Avaandmete rohelist raamatut ei ole uuendatud peale selle valitsuse poolset vastuvõtmist aastal 2014. Avaandmete roheline raamat on valdkonna poliitika põhimõtteid avav dokument, kuid ei reguleeri sisu täpsemalt ega anna konkreetseid juhiseid näiteks selles osas, mis formaadis avaandmeid peaks avaldama. Avaandmete roheline raamat annab soovitusel andmehulga avaldamisel lähtuda Tim Berners-Lee tärnide süsteemist [17]. Avaandmete rohelises raamatus täpsustatakse, et sobivad vähemalt kolmas ja kõrgem taseme. See tähendab, et andmed peavad olema masinloetaval kujul ning formaat ei või olla seotud ühegi konkreetse tootega, vaid vabavaraliste lahendustega töödeldav. Selliselt on töödeldavad näiteks ODF-formaadis dokumendifailid ja struktuursete andmete puhul levinuimad formaadid CSV, JSON, XML [5].

Avaandmete portaalis on kasutusjuhend andmete avaldamiseks ja uuendamiseks. Samuti on juhised andmete metaandmete portaali lisamiseks ehk informatsiooni, mis on portaali andmehulga avakavas lisainformatsioonina nähtav.

2.5 Avaandmete mõjuhinna

Irimaal on avaandmete avaldamise toetamiseks loodud andmete auditi tööriist, mida kõik saavad kasutada [18]. Tegemist on abistava tööriistaga andmete omanikule, kes saab auditi läbi tegemisel teada, kas andmeid avaldada ning kuidas. Irimaa loodud auditi tööriist on toetavaks rakenduseks teabevaldajatele ning selle sisu ja eesmärgi võib võrrelda AvTS tuleneva mõjuhinna kohustusega. Irimaa küsimustiku eesmärk on ennetada, et avaldatakse avaandmeid, millel pole väärtust ehk huvi nende kasutamise osas puudub või ei ole teada.

⁵ Avaandmete loomise ja avaldamise juhend: https://www.ria.ee/sites/default/files/content-editors/publikatsioonid/avaandmete_loomise_juhend.pdf

Irimaal on avaandmete auditi eesmärgiks ka võimaldada organisatsioonil saada ülevaadet kõikidest andmekogumitest ehk see on toetav töövahend. Eestis on andmekogudest ülevaate saamiseks olemas Riigi infosüsteemi haldussüsteem (RIHA), kus on riigi infosüsteemi terviklik ja üksikasjalik kaardistus. RIHA kaudu saab ülevaate, millised infosüsteemid ja andmekogud on liidetud riigi infosüsteemiga, milliseid andmeid kogutakse, milliseid andmeteenuseid osutatakse. Kogu riigi infosüsteemi hallatakse RIHA kaudu [19]. RIHA-s menetletakse infosüsteemide ja andmekogude asutamist, registreerimist, liitumist X-teega ja teisi riigi infosüsteemi eluks vajalikke protsesse [19]. Seega puutub iga andmekogu või infosüsteemi looja ning pidaja kokku RIHA-ga. RIHA reguleerib Riigi infosüsteemi haldussüsteemi määrus⁶. Kõik andmekogud, mis on registreeritud RIHA-s, ei kuulu avaandmete hulka, aga ülevaate saamiseks, mis andmeid võiks keegi mõne rakenduse jaoks vajada, peaks RIHA-st saama.

Isikuandmeid sisaldavaid andmeid on võimalik avaandmetena avaldada, kui on läbi viidud AvTS § 3¹ ettenähtud mõju hindamine. Mõjuhindang tuleb teha, kui isikuandmete töötlemise laad, ulatus, kontekst või eesmärk võib kaasa tuua inimestele suure füüsilise, varalise või mittevaralise kahju ehk ohu ning andmetöötlus on ulatuslik ja süsteemne [20]. Mõjude hindamise kohustus on uus ning järeldused teeb läbiviija ise, mistõttu ei pruugi see olla alati täpne ning asutused võivad vajada lisaks juhendmaterjalidele täiendavat tuge. Sellest tulenevalt on magistr töö üks eesmärk pakkuda praktilise lahendusena vahend, mis aitaks teabevaldajatel avaandmete mõjuhindangut läbi viia.

Andmekogud, mis sisaldavad isikuandmeid saab anonümiseerida ja pseudonümiseerida, et eemaldada konkreetse isiku kohta käivad. Pseudonümiseerimisel isiku tuvastamist võimaldavad andmekirje väljad asendatakse ühe või mitme väljamõeldud identifikaatoriga [21]. Ükskõik, kui hästi krüpteeritud pseudonüüm on loodud, ei taga pseudonümiseerimine täielikult isiku kaitset [22]. Anonümiseerimine on protsess, mille käigus eemaldatakse tunnused, mille abil on võimalik isikut iseloomustada ja/või tuvastada [22]. Kui andmetest on eemaldatud konkreetset isikut iseloomustav info, siis ei saa andmeid enam nimetada isikuandmeteks. Kui asutused järgivad ja täidavad GDPR⁷ nõudeid, soodustab see avaandmete avaldamist ja puuduvad takistused nende andmete taaskasutamiseks. See omakorda soodustab andmepõhist majandust ning võimaldab uute toodete ja teenuste loomist [23].

2.6 Rahvusvaheline praktika avaandmete valdkonna toetamiseks

Teiste riikide kogemusest õppimiseks on oluline tutvuda nende avaandmete valdkonna praktikaga ja infomaterjalidega. Lähtuvalt sellest, et materjale ja riike mille kohta infot otsida on palju, tegi autor valiku sissejuhatuses viidatud edetabelite tippudes olevate riikide hulgast. Maailma parimate avaandmete riikide kohta on eraldi koostatud 2018. aastal aruanne *Open Data Barometer: Leaders Edition, From Promise to Progress*. Avaandmete valdkonna maailma edukamad riigid on Kanada, Suurbritannia, Austraalia, Prantsusmaa ja Uus-Meremaa. Maailma edetabelis jagavad avaandmete vallas esikohta Iirimaa ja Kanada, mille kohta on käesolevas magistr töö täpsem ülevaade. *Open Data Barometer: Leaders Edition, From Promise to Progress* soovib valitsustel sätestada konkreetset juhised, tehnilised standardid ja/või menetlusprotsess, kuidas avalik sektor peab andmeid avaldama, et see toimuks kõikides asutustes ühtsetel põhimõtetel [24]. See tagab andmete lihtsama kasutamise ja mõistmise.

Irimaal on valdkonna arenguks eraldi koostatud avaandmete strateegia, mis on täienduseks üldisele IKT strateegiale [25]. Eestil on valdkonnas üks katus-strateegia Infoühiskonna arengukava 2020, mis põgusalt käsitleb ka avaandmete valdkonda. Iirimaa strateegia erineb Eesti valdkonna strateegiast ka seetõttu, et seal on kajastatud reaalsed näited erinevate

⁶ Riigi infosüsteemi haldussüsteemi määrus, <https://www.riigiteataja.ee/akt/13147268?leiaKehtiv>

⁷ GDPR – General Data Protection Regulation ehk isikuandmete kaitse üldmäärus

valdkondade avaandmete kasust ehk info erinevatest rakendustest, mis on avaandmete abil loodud ja kasutusel. Irimaa avaandmete strateegias on 7 suurt eesmärki, mille saavutamiseks on kirjeldatud mitmed alamtegevused [25]. Ühe eesmärgina on plaanis laiendada avaandmete ja nendega seotud metaandmete ulatust ning parandada andmete kvaliteeti, kvantiteeti ja valikut ehk mitmekesistada valikut riiklikus avaandmete portaalis [25].

Irimaal on loodud tehniline raamistik, mille eesmärk on abistada avaandmete avaldamise protsessi tehniliselt poolelt. Irimaal on kohustus avaandmeid avaldada Tim Berners-Lee [17] tärnide süsteemi järgides vähemalt kolmandal tasemel [26]. Tehniline raamistik sisaldab ka juhiseid lähtuvalt andmete valdkonnast, et mis formaadis peaks andmeid avaldama, kui tegemist on näiteks üldiste andmetega või kui tegemist on ruumiliste andmetega [26]. Loodud on standard avaandmete metaandmete avaldamiseks, mida on soovituslik järgida [26]. See annab andmete kasutajale ka selguse ja ülevaate andmekogu sisust. Eestis on soovitusel erinevate metaandmete standardite järgimiseks ja kasutamiseks andmete kohta käivate andmete kirjeldamisel, kuid kohustust pole. See teeb omakorda keerukamaks andmete kasutamise huvitatud osapooltele.

Kanada on tõusnud juhtivaks riigiks avatud valitsemise ja avaandmete valdkonnas. Lisaks föderaalvalitsusele, on Kanadas paljud provintsid ja kohalikud omavalitsused käivitanud avaandmete algatusi. Kanada on algatanud ka Vaikimisi Avatud (*open by default*) initsiatiivi, eesmärgiga avaldada kõiki kogutud andmeid, et toetada valitsuse püüdlusi läbipaistvusele [27].

Kohalikele omavalitsustele ja riigiasutustele on loodud „tööriistakast“ (*DIY Open Data Toolkit*), mis sisaldab erinevat materjali toetamiseks ja juhendamaks andmete avalikustamist [28]. See koondab endas koolitusmaterjale, parimaid praktikaid, tööriistu ja teisi abinõusid, mis aitavad ette valmistada andmete avaldamist [28]. Eesmärk ei ole ainult uute süsteemide loomisel juba andmete avaldamise poliitikat juurutada, vaid aitab olemasolevaid andmehulki teha avalikkusele kättesaadavaks [28]. Tegevus on jagatud mitmesse etappi, avaandmete valdkonnaga tutvumisest kuni andmete täieliku avaldamiseni [28]. „Tööriistakasti“ on kokku kogutud edumeelsemate omavalitsuste kogemused ja parimad praktikad, mille eesmärk on jagada teiste kogemusi ja kirjeldada tegevuskava andmete täieliku avaldamiseni [28].

2.7 Arendusprojektide toetusmeetmed

Avalikule sektorile on avaliku teabe seadusega pandud kohustus avaandmeid avaldada. Sealjuures on tagatud ka rahaline toetusmeede kohustuse täitmiseks. Perioodil 2014–2020 toetab Euroopa Liit infoühiskonna arengut Eestis 181 760 000⁸ euroga, millele lisandub Eesti riigi kaasfinantseering 24 928 235 eurot. Infosüsteemide arendamiseks on eraldatud sellest summast 93 815 944 eurot [29]. Toetuse saamiseks on vaja esitada taotlus RIA-sse. Toetuse andmist reguleerib majandus- ja taristuministri määrus nr 31 „Avalike teenuste pakkumise arendamiseks toetuse andmise tingimused ja kord“. Määruse alusel antakse erinevaid toetusi, mis on seotud infotehnoloogiliste arendustega ning meede on omakorda jagatud erinevateks voorudeks. Eraldi on toetusvoor „Avaandmete ja andmeanalüüsi kasutuselevõtu edendamine“, mille otsene eesmärk on toetada avaandmete avaldamist. Taotlusvooru eesmärk on andmeanalüüsi meetodite ja andmete reaajas jälgimise võimaluste abil luua eeldused avalike teenuste kvaliteedi tõstmiseks ja paremate juhtimisotsuste tegemiseks. Taotlusvooru tulemusel on paranenud avaliku sektori andmete kättesaadavus, suurenenud avaandmete kasutuselevõtt nii avalikus kui ka erasektoris ning avaliku sektori otsuste tõenduspõhisus on kasvanud tänu

⁸ Ühtekuuluvuspoliitika fondide rakenduskava 2014-2020,
https://www.struktuurifondid.ee/sites/default/files/uhtekuuluvuspoliitika_fondide_rakenduskava_2014-2020_211218.pdf

andmeanalüüsi kasutamisele [30]. Taotlusi on saanud esitada alates 30. maist 2016. Taotluste vastuvõtmine lõpetatakse, kui esitatud taotluste kogumaht ulatub 5 000 000 euroni. 2019. aasta 30. märtsi seisuga on lõppenud kaks projekti, kogumaksumusega 154 684,8 eurot⁹. Avaliku sektori asutustele on seadusega antud ülesande täitmiseks ja avaandmete avaldamiseks võimaldatud toetust taotleda, kuid ka see pole valdkonnas aidanud.

Nõuded projekti taotlusele on reguleeritud majandus- ja taristuministri määrusega „Avalike teenuste pakkumise arendamiseks toetuse andmise tingimused ja kord“. Lisaks määrusele on juhend, mis selgitab põhjalikult kõiki projekti taotluse osasid¹⁰ ning selgitab, mis infot lugeja vastavas punktis soovib näha. Sealhulgas on üheks kohustuslikuks punktiks lahter „Projekti eeltingimused“. Selle punkti raames peab taotleja selgitama, kuidas ta suhestub RIHA¹¹, ISKE¹², koosvõimeraamistiku¹³, iseteeninduse raamistiku¹⁴ ja kasutatavuse raamistikuga¹⁵ ning kas arendus vastab riigipilve alusnõuetele¹⁶. Lisaks peab taotleja projekti eeltingimustes kirjeldama, mis on avaandmete hetkeseis ning mis on plaanid avaandmetega seoses planeeritava arendusega. Vastavalt juhendile tuleb täpsustada, kas loodavale arendusele laieneb AvTS-ist tulenev kohustus andmete masinloetavas formaadis kättesaadavaks tegemise osas. Juhul, kui AvTS-i järgi vastav kohustus kehtib ja/või tekib, siis tuleb selgitada kohustuse täitmise hetkeseisu ning projekti käigus planeeritavaid tegevusi nõuete täitmise osas. See tähendab, et tuleb täpsustada, millisel kujul, mis ajaks ja kuhu andmed avalikult kättesaadavaks tehakse.

⁹ Jooksev voor: avaandmete ja andmeanalüüsi kasutuselevõtu edendamine, <https://www.ria.ee/et/struktuuritoetused/2014-2020/avalikud-teenused/voor-avaandmed.html#projektid>

¹⁰ Täpsemalt saab lugeda siit: <https://www.ria.ee/sites/default/files/content-editors/SF/juhend-taotluse-vormistamiseks-2018.pdf>

¹¹ <https://www.riha.ee/Avaleht>

¹² <https://www.ria.ee/et/kuberturvalisus/iske/juhendid-ja-materjalid.html>

¹³ https://www.mkm.ee/sites/default/files/iseteeninduskeskkondade_raamistik.pdf

¹⁴ https://www.mkm.ee/sites/default/files/kasutatavuse_moodikute_susteem_final_1.pdf

¹⁵ https://www.mkm.ee/sites/default/files/kasutatavuse_moodikute_susteem_final_1.pdf

¹⁶ https://www.ria.ee/riigiarhitektuur/wiki/lib/exe/fetch.php?media=an:riigipilve_alusno_uded_v.2.pdf

3 Metoodika

Magistritöö uurimuse läbiviimiseks on valitud kvalitatiivne meetod, mis võimaldab uuritava valdkonna tundma õppimist paindlikult. Kvalitatiivse lähenemise eesmärgiks on valdkonna mõistmine ja tõlgendamine. Kvalitatiivse meetodi puhul ei ole sihiks uuritava valdkonna kohta üldistavate järelduste tegemine, vaid selle mõtestamine ja järelduste tegemine kogutud informatsiooni, analüüsi põhjal. Denzin ja Lincoln selgitavad, et kvalitatiivne uuring hõlmab endas tõlgenduslikku ja naturalistlikku lähenemist maailmale [31]. See tähendab, et kvalitatiivsetes uuringutes uurijad uurivad nähtusi nende loomulikus keskkonnas, püüdes ilminguid mõtestada või tõlgendada nende tähenduste kaudu, mida uuritavad annavad [31]. Creswell toob välja, et kvalitatiivses uuringus kasutab uurija kompleksset järeldusteni jõudmise viisi [32]. Creswell leiab, et järeldusteni jõudmine on suuresti induktiivne, kuid töös on nii induktiivsed kui ka deduktiivsed protsessid [32]. Mõtlemisprotsess on interaktiivne ehk kordumistele toetuv – andmekogumiselt ja –analüüsilt liigutakse tsükliliselt uurimisküsimuste juurde ja sealt jälle tagasi. Andmete kogumist, analüüsimist ja sellest kirjutamist võib mõista kui üheaegseid tegevusi.

Avaandmete valdkond pole Eestis võrreldes muu digitaalse arenguga sama kiirelt kaasa läinud. Avaandmeid ei avaldata meelsasti ning teisalt pole teada ka huvi nende kasutamise osas. Lähtuvalt sissejuhatuses kirjeldatud olukorrast on magistritöö eesmärk pakkuda välja avaandmete mõjuhinnangu töövahend. Selleks, et praktiline töövahend sisustada ning sellele spetsialistide tagasiside küsida, on oluline Eesti avaandmete valdkonda ja hetke olukorda mõista. Selleks viib töö autor läbi projektide analüüsi ning pool-struktureeritud intervjuud avaliku sektori avaandmete valdkonna spetsialistidega. Intervjuud on plaanis teha vähemalt ühe inimesega, kes tegelevad avaandmete valdkonnaga, igas valitsemisala organisatsioonis.

Magistritöö eesmärgi saavutamiseks on sõnastatud kolm uurimisküsimust:

- 1) Miks on avaldatud avaandmete maht väike?
- 2) Kuidas edendada avaandmete avaldamist?
- 3) Kas ja kuidas avaandmete mõjuhinnangu koostamise vahend toetab andmete avaldamist?

3.1 Andmekogumise meetod

Avaandmete avaldamisest ülevaate saamiseks on üheks töö analüüsi osaks võetud Euroopa struktuurivahendite perioodi 2014-2020 jooksul RIA-sse esitatud infosüsteemide arendusprojektid. Projektidele ligipääs on autoril töökohustuste täitmise tõttu. Käesoleva magistritöö analüüsi tegemise aluseks on kõik arendusprojektid alates taotlusvoorust avanemisest, mis on esitatud seisuga 10.02.2019 rahastamise taotlemiseks. Kokku on esitatud perioodi vältel taotlusi, mis on saanud positiivse otsuse, 169. Projektidel on erinevad staatused. Kõikidest projektidest 64 on lõppenud projektid. Ülejäänud 105 projekti on saanud positiivse otsuse ning on arendamise etapis. Lõppenud projektid on need, mis on valmis arendatud pakkuja poolt, kes on läbi hankemenetlusprotsessi leitud. Hankes tellitud ja kirjeldatud soovitud funktsionaalsus on loodud ja töökorras ning RIA poolt kontrollitud. Lõppenud staatusega projektid on kasutuses tervikliku infosüsteemina või selle osana.

Lisaks projektidele ja nende analüüsile, kogutakse valdkonna spetsialistidelt andmeid pool-struktureeritud intervjuude kaudu. Töö autoril oli eesmärk teha näost-näku intervjuud vähemalt ühe spetsialistiga igast valitsemisalast, mis on avaandmeid avaldanud. Kaitseministeerium ei ole avaandmeid avaldanud ning nendega ei võtnud autor kontakti. Autor võttis ühendust 12 inimesega, kellest 11 sai positiivse vastuse intervjuul osalemiseks. Ühest

valitsemisalast ei saanud autor tagasisidet ning seetõttu ei ole ühte valdkonda uurimusse kaasatud. Autor tegi kokku 11 intervjuud 14 inimesega (kolmel intervjuul osales korraga kaks inimest). Osad intervjuud on magistritöö raames tehtud tehnilise taustaga inimestega ning teised äripoolse vastutavate spetsialistidega. Intervjueeritavad oli ametinimetuste poolest infotehnoloogia osakonna juhataja, analüütik, infotehnoloogia osakonna peaspetsialist, infotehnoloogia osakonna nõunik, arendusjuht, projektijuht, arhitekt, peaspetsialist (äripoolse esindajana ehk andmekogu peakasutaja roll). Olenemata organisatsioonist, valdkonnast või ametinimetusest olid kõik intervjueeritavad seotud avaandmete temaatikaga ning organisatsioonid oli avaandmeid avaldanud.

Lisaks intervjueeriti valdkonnaga tegelevaid inimesi Andmekaitse Inspeksioonist ja Majandus- ja Kommunikatsiooniministeeriumist poliitikakujundamise eest vastutavat inimest. Mõjude hindamise töövahendile on täiendavalt tagasiside küsitud kahelt avaandmete valdkonna ekspertidelt.

Kõige pikem intervjuu kestis 94 minutit, kõige lühem intervjuu kestis 36 minutit. Intervjuud kestsid keskmiselt ca 58 minutit. Esimesed kuus intervjuud autor salvestas ning transkribeeris intervjuu järgselt. Järgmisi intervjuusid autor ei salvestanud, sest osaliselt hakkas intervjuude käigus esitatud informatsioon korduma. Edaspidi tegi autor intervjuude käigus märkmeid. Märkmete tegemiseks oli ette olemas varasemad kodeerimise märksõnad, mille taha autor lisas vastavalt pluss märgi, kui sama sisu, samade märksõnadega intervjuul teemaks tuli. Uut vestluselt saadud infot lisas autor vaba tekstina, mida hiljem analüüsis kasutas.

Intervjuude läbiviimiseks oli koostatud pool-struktureeritud kava, mis oli disainitud peamiselt avatud küsimustest ja millele lisaks esitas autor täiendavaid küsimusi lähtuvalt vestluse käigust. Koostatud kava lähtus peamiselt uurimisküsimustest ning oli jaotatud tinglikult neljaks osaks. Intervjuu sisaldas üldist osa, mille raames uuriti organisatsiooni lähenemist avaandmete temaatikale, milline on valmisolek andmete avaldamisele ning missugune on teadaolev huvi valdkonna avaandmete osas. Üldisest osast edasi olid küsimused konkreetsemalt valdkonna avaandmete kohta – kuidas andmeid avaldatakse, mitmele tähtsusele vastavad avaldatud avaandmed, kuidas on uute arendusprojektide puhul lahendatud avaandmete temaatika, kuidas on avaandmete avaldamise protsess korraldatud, kuidas on tagatud isikuandmete kaitse, turvalisus. Viimane osa intervjuust käsitles soovitusi ja ettepanekuid, kuidas Eestis avaandmete valdkonda edasi arendada. Täpne intervjuu kava on leitav käesoleva töö lisast I.

Intervjuudel tutvustati avaandmete avaldamise mõjude hindamise töövahendit. Eesmärk oli töövahendile saada tagasiside kahes osas. Üks pool oli seotud konkreetsete küsimustega ehk kas need aitavad teabevaldajal teha otsust andmete avaldamise osas või mitte. Teisalt oli soov saada tagasiside, kas antud töövahendit peetakse vajalikuks ja abistavaks meetmeks või mitte.

Lisaks projektidele ja intervjuudele tutvus autor valdkonnast ülevaate saamiseks valitsemisala IKT strateegiatega, avaandmete portaalis avaldatud infoga ja valdkonna juhendmaterjalidega, mis aitasid konteksti paremini mõista ja valdkonda tundma õppida. Samuti oli oluline roll teooria, artiklite ja teiste uurimuste lugemisel. Lisaks kontakteerus autor Iirimaa avaandmete organisatsiooniga (*Open Data Unit, Department of Public Expenditure and Reform*), et paremini mõista nende *Open Data Audit Tool*'i [16] ja selle eesmärke, mis on osaliselt aluseks võetud käesoleva töö avaandmete avaldamise mõjude hindamise töövahendi loomisel.

3.2 Andmete analüüsi meetod

Projektide ja intervjuude käigus kogutud andmete analüüsimiseks on kasutatud kvalitatiivse sisuanalüüsi metoodikat. Kvalitatiivne sisuanalüüs võimaldab analüüsimisel paindlikkust [33]. Andmete analüüsimiseks on kasutatud intervjuude transkriptsioone ja intervjuude käigus tehtud

autori märkmeid. Kvalitatiivse sisuanalüüsi puhul ei ole alati eesmärgiks uuritavat teksti analüüsiühikute kaupa kodeerida, kuid kodeerimine on lubatud ja analüüsi protsessi käigus on võimalik koode lisada, eemaldada, täiendada [34]. Analüüsi jaoks kasutas autor intervjuude transkriptsioone ja märkmeid. Intervjuude ja nende transkriptsioonide analüüsimisel kasutati avatud kodeerimist, kus materjalide korduval lugemisel ja oluliste märksõnade markeerimise tulemusena selguvad koodid [35]. Kodeerimise käigus tekkisid sõnad, laused, mille autor markeeris teise tooniga ja koondas kokku ühe koondteema alla. Intervjuudeks, mida ei salvestatud, oli autor enne intervjuud märksõnad ette kirjutanud ning tegi vestluse käigus nende teha lisamärkmeid. Juhul kui intervjuul selgus uus info, mida varasemalt polnud kodeeritud, kirjutas ta selle tekstina üles ja lisas kodeeringu osaks, mida hiljem analüüsis. Kodeerimise tulemuste põhjal lõi autor suuremad kategooriad, kuhu alla koondas erinevate intervjuude samasisulise info. Loodud kategooriatesse koondatud sisu alusel viis autor läbi analüüsi. Autor luges analüüsi protsessi käigus korduvalt läbi intervjuude transkriptsioone ja märkmeid.

Kvalitatiivseid meetodikaid kasutades pole magistritöö eesmärgiks tulemuste üldistamine, vaid ülevaate saamine avaandmete valdkonnast kogutud info põhjal ja selle edasi andmine töö lugejale.

4 Projektide ja intervjuude analüüs

Järgmises peatükis antakse ülevaade projektide ja intervjuude analüüsi tulemustest. Peatükk on jagatud kaheks osaks, millest esimene pool keskendub projektidele ja teine intervjuude tulemuste analüüsimisele. Analüüs annab ülevaate avaandmete valdkonna praegusest olukorrast ning selgitab, mis on avaandmete avaldamist takistavad tegurid.

4.1 Projektide analüüsi tulemus

Intervjuudel saadud info hakkas osaliselt korduma pärast viiendat vestlust. Intervjuud andsid osaliselt sisendi, mida oleks vaja täiendavalt teha ehk kuidas saaks paremini toetada avaandmete valdkonna arengut riigis. Intervjuudest tulid välja ka põhjused, miks seni on avaandmete avaldamise kohustusse suhtunud pigem leigelt.

Projektide analüüs jaotati mitmeks etapiks. Esmalt koostati nimekiri projektidest, mis käesoleva magistr töö uurimise eesmärgiks sobivad. See tähendab, et kogu Euroopa Struktuurifondi perioodi 2014-2020 projektitaotlustest¹⁷ tehti valik, vaid neile projektidele, mis on antud töö eesmärgist lähtuvalt asjakohased ehk tegemist on infosüsteemide arendusprojektidega. Magistr töö uurimise jaoks sobivaid ja lõppenud arendusprojekte oli 64. Need projektitaotlused luges autor läbi ning koondas Excelisse. Projektide nimekirjale lisati projekti nimi, taotleja organisatsioon, projekti staatus, eeltingimuste lahtri kirjeldus, projekti lühikirjeldus, avaandmete olukord, autori märkused. Viimased kaks lahtrit on autori enda märkmeteks analüüsi koostamisel. Peale esmast projektide läbitöötamist jäi alles 34 arendusprojekti, milles oli infot tekkivatest avaandmetest ning nende avaldamisest. Analüüsi teine etapp seisnes kontrollis, kas antud lubadusi on täidetud. Selleks uuris autor riiklikku avaandmete portaali või teisi kodulehekülgi, millele oli projektides viidatud. Osad avaliku sektori asutused on loonud enda veebilehele avaandmete lingi, millelt saab infot. Nende puhul, kes olid nõudeid täitnud ja avaandmed avaldanud, vaatas autor analüüsi kolmandas etapis erinevaid aspekte nagu näiteks metaandmete kirjelduse avaldamine, seda mis formaadis on avaandmed avaldatud. Samuti vaatas autor märgitud litsentse ning seda missugune on andmete uuendamise sagedus.

Järgnev analüüs puudutab eelpool nimetatud väikest osa avalikus sektoris tehtud arendusprojekte ja avaandmeid, mis on sisendiks magistr töö järgmiseks osaks. Esmane analüüs avaandmete avaldamisest annab sisendi intervjuude ettevalmistamiseks.

Erinevaid andmehulki riiklikus avaandmete portaalis läbi lugedes jääb esmalt silma, et andmete kohta on olemas üldjuhul mõistetav ja selge lühikirjeldus ehk lihtsasti on aru saada, mis valdkonna andmetega on tegemist. Samas ei ole etteantud juhust või põhimõtteid, mida üldises kirjelduses peaks kajastama. Metaandmete kirjeldamine ja avaldamine on kõigil erinevalt lahendatud, kui üldse. Kui andmed avada, saab järeldada, mis andmeväljade tähendused on, kuid masinloetavas formaadis avaandmete kohta käivaid metaandmeid oli vähe. Kirjeldamata on ka üksikute andmehulkade struktuur.

Avaandmete portaali andmehulga avakuval on näha, et andmetele on määratud *Creative Commons* 3.0 litsents. See tuleneb sellest, et aasta alguses otsustasid avaandmete portaali haldajad, et nendel andmehulkadel, millele ei ole määratud litsentsi, lisatakse portaali haldajate poolt eelpool nimetatud litsents. Seda otsustati põhjusel, et kui andmetel ei ole määratud näiteks *Creative Commons* 3.0 või muud samaväärset litsentsi, siis pole kasutajatel kindlust, et neid avaandmeid võib kasutada ka näiteks ärilistel eesmärkidel. Samas võib see lähenemine tekitada andmetest huvitatud osapooltes segadust, sest riiklik avaandmete portaal on suures osas lingikogum, mis viitab avaandmete tegelikule asukohale. Kui kasutaja avab lingi ning läheb

¹⁷ Lisaks on projektid, mille eesmärgiks on analüüsi läbiviimine ja/või baasinfrastruktuuri uuendamine.

edasi lehele, kust ta saab teha soovitud andmetele väljavõtte, siis uuel lehel viidet või märget litsentsi osas pole.

Andmed on avaldatud erinevates formaatides. Suur osa andmeid on esitatud andmetömmistena registritest, andmebaasidest. Projektide hulgas oli API¹⁸ lahendusi üksikutel andmehulkadel, kuid laiemalt on avaandmete portaalis viidatud andmed kättesaadavad ka API-de kaudu. Analüüsi käigus selgus, et avaldatavad andmetömmised võetakse osaliselt otse andmebaasist ilma, et enne avaldamist tehtaks isikuandmetele anonümiseerimist. Võimalik on alla laadida Exceli tabelina teatud teenuse tarbijad koos nimede, aadresside ja erinevate kontaktandmetega. Saab öelda, et probleemiks on andmete avaldamise oskus. Andmete anonümiseerimise ja/või andmeväljade avaldamata jätmise protsess peaks olema paigas enne avaldamist.

Andmete juures puudub info nende uuendamissagedusest. Andmed on suuresti kättesaadavad allalaadimise teel tabelina ning seal puudub nende kogumise kohta kuupäev või periood. Andmete kasutajal on keeruline nende andmete pealt teha edasisi otsuseid.

Rohelises raamatus ja Eesti avaandmete poliitikas on lähtekohaks, et masinloetavuse nõude täitmiseks on avaandmed esitatud vähemalt kolmandale tärnile vastavas formaadis: nende kasutus on määratud vaba litsentsiga, neid saavad kasutada mistahes tarkvararakendused, nad järgivad avatud standardeid [5]. Analüüsi käigus selgus, et osaliselt olid asutused lubanud kõrgemale tärnile vastavalt andmeid avaldada, kuid tegelikkuses seda ei tehtud. Kuigi kolmandale tärnile vastavalt andmete avaldamine on soovituslik, siis ettepanek on, et seda peaksid avaliku sektori asutused järgima.

Analüüs lõppenud projektidest annab arusaamise, et valdkonna arenguks on võimekust. Riiklikus avaandmete portaalis on 155¹⁹ andmekogu. Enne projektide läbivaatamist oli neid ca 80. Pärast teostatud kontrolli võeti kõikide projektide vastutajatega ühendust ning küsiti selgitusi, miks pole lubatud teostatud. Seejärel lepiti kokku ajas, mille jooksul lubatud andmed avaldatakse. Suurem osa organisatsioonide on tagantjärele lubadused täitnud.

4.2 Intervjuude analüüsi tulemus

Intervjuude tulemuste analüüs on esitatud alapeatükkide kaupa, mis on peamised intervjuude analüüsis selgunud teemad.

4.2.1 Avaandmete avaldamise kohustus

AvTS kohustab masinloetavalt avaldama vaid teatud osa andmetest. Osadel ministriumidel on lähenemine, et avaandmetena avaldatakse samu andmeid, mis on kättesaadavad ka läbi avalike registripäringute. Kuigi läbi registrite andmeid pärides võib sealt leida näiteks inimeste nimesid, siis avaandmetena on avaldatud anonümiseeritud andmed ja nimesid ei avaldata. Intervjuudel avaldasid spetsialistid arvamust, et avalike registrite kaudu kasutatakse andmeid oluliselt rohkem kui samu andmeid avaandmetena. Mitmel spetsialistil oli arvamus, et kodanikel, ajakirjanikel on huvi leida infot ja lugeda materjale, mitte andmeid kasutada masinloetavalt. Intervjuudest koorus ettepanekuna idee, et kuigi info on avalikult kättesaadav, siis kõike ei pea tegema masinloetavaks kättesaadavaks. Avatuse tagamiseks piisab ka sellest, kui andmed on mitte masinloetavalt avaldatud.

Analüüsi käigus selgus, et andmekogude avaandmetena kättesaadavaks tegemisega on hästi. Andmekogude osas on põhimäärused, mis ütlevad, mis andmeid avalikustatakse. Juhul, kui andmekogusse lisatakse andmevälju, aga põhimäärust ei uuendata, võib juhtuda, et kõikvõimalikke andmeid pole avaldatud. Kuna avaandmete valdkond pole igas organisatsioonis

¹⁸ *Application Programming Interface* ehk rakendusliides

¹⁹ Avaandmete portaali andmekogude ülevaade 28.03.2019 seisuga

prioriteediks, siis regulaarseid hindamisi või analüüse valdkonnas ei tehta ega kontrollita, kas kõik võimalikud andmed on avaandmetena avaldatud. Samuti puudub riiklik järelevalve, kes nõude täitmist kontrolliks.

Avaandmete avaldamist mõjutab see, et kõikide arendatud infosüsteemide puhul ei ole nende loomise hetkel arvestatud avaandmete avaldamise kohustusega. See on suureks takistuseks mitmes valdkonnas. Võimekuse loomist peetakse kalliks ja seetõttu pigem loobutakse ning loodetakse, et teemaga ei pea tegelema. Uute arenduste planeerimisel arvestatakse koheselt avaandmete avaldamise nõudega. Sellega seoses vaadatakse üle planeeritav andmekoosseis ja see omakorda arutatakse läbi juristidega, kes aitavad otsustada, missuguseid andmeid võib avaandmetena kättesaadavaks teha. Kui varasemalt kiputi avaandmeid unustama, siis täna enam arenduste käigus seda ei juhtu. Mitmel korral rõhutatakse intervjuude käigus, et avaandmed on rahastuse taotlemise eeltingimuseks ja seda peab tegema.

4.2.2 Huvi avaandmete kasutamise osas

Igal intervjuul tuli teemaks avaandmete kasutamine ja huvi nende osas. Täna ei ole teada, milline on avaldatud avaandmete kasutamine. Selle üle ei toimu monitooringut. Üks intervjuueeritav oskas tuua statistilisi numbreid avaandmete kasutamise kohta, teised seda infot intervjuudel ei maininud ja/või seda ei jälgita. Avaandmete portaali haldajad teavad, mis on portaali külastatavus, kuid see pole samaväärne või võrreldav andmete kasutamisega sellel eesmärgil, milleks on nad masinloetavalt avaldatud. Mitmed inimesed ütlesid intervjuude käigus, et esmalt on oluline teada saada, kas on huvi andmete järgi või mitte. Avaandmete avaldamise protsessile paratamatult kulub raha ning kui huvi ja reaalne kasutamine puudub, siis kas tehtavad kulutused on eesmärgipärased ja vajalikud. Mitmel intervjuul tuli välja, et kui keegi küsib andmeid, mille avaldamiseks puuduvad piiranguid, siis ollakse valmis alati andmeid huvilistele kättesaadavaks tegema ning seda ka masinloetavalt. Teisalt teabevaldajad ka ise keskselt ei jälgi, mis andmete vastu huvi tuntakse. Osaliselt on teabevaldajatele tehtud pöördumisi avaandmete osas ning on teada kasutuslood. Avaandmete portaali kaudu saab esitada andme soove, mis esitatakse läbi Githubi²⁰ vastava valdkonna avaandmete spetsialistile. Otse teabevaldajate poole pole andmesoovidega pöördutud, kuid seda eelistaksid mitmed. Soov on avaldada masinloetavalt neid andmeid, mida kasutatakse. Tehti ettepanek, et lähenemine avaandmete avaldamise osas võiks olla vajaduspõhine, et mitte ressursse raisata. See tähendaks, et avaandmed ei pea kogu aeg reaalajas olema kättesaadavad, kui huvi nende järgi pole, vaid need tehakse kättesaadavaks, kui keegi vastavasisulise päringu teeb. Mitmed inimesed ütlesid vestluste käigus, et kui oleks teada, et avaandmete järgi on reaalne vajadus, siis on valmisolek sellesse panustada. Kuna täna ei mõisteta ega nähta suuresti väärtust avaandmete avaldamisel, siis ei soovita tööaega sellesse panustada.

Analüüsi käigus selgus, et osaliselt on avaandmete mitte avaldamise põhjuseks see, et ei nähta selle tegevuse mõju. Arvatakse, et avaandmeid ei kasutata rakenduste loomiseks ning seetõttu võib jääda arusaam, et tegelikult pole ka neid kellelegi vaja. Samas ei ole täna teada, kellel on huvi või mitte, sest seda infot otseselt organisatsioonid ei koonda, et mis asutus on andmete saamiseks nende poole pöördunud ning mis eesmärgil soovitakse andmeid kasutada. Avaandmete portaali võiks olla avaldatud statistiline info, mis andmeid ja kui palju kasutatakse otse masinloetavalt ehk päringuid tehakse API kaudu liidestatud rakendustest.

Arvati, et kui rohkem avaandmete teemal teadlikkust tõsta ning teavitada tehtud rakendustest, siis on ka teabevaldajatel rohkem huvi andmeid avaldada. Kui on teada reaalsed kasutuslood, siis see innustab andmeid avaandmetena kättesaadavaks tegema. Avaandmete portaalis on kirjas osad rakendused, mida on loodud, kuid arvatavasti pole see kogumaht. Avaandmete valdkonna arendamise kasu on vaja selgitada läbi reaalsete näidete. Teadlikkust saab tõsta ka

²⁰ Ülevaade küsitud andmetest ja üleval olevatest päringutest: <https://github.com/okestonia/opendata-issue-tracker/issues>

läbi valitsemisalade allasutuste. Ministeeriumi IT-juhil on võimalus valdkonna arengusse panustada ka läbi hallatavate asutuste teavitamise. Näiteks on võimalik edastada infot, aga ka küsida allasutustelt tagasiside avaandmete kohta lähtuvalt AvTS kohustusest. Seeläbi hakkavad ka väiksemad organisatsioonid teemale mõtlema ja võibolla leiavad, mida oleks vaja avaandmetena avaldada.

4.2.3 Avaandmete avaldajad

Intervjuude käigus juhiti tähelepanu sellele, kes võiksid avaandmeid avalikustada. Riik on laiem, kui ministeeriumid ning nende hallatavad andmekogud. Avalikke ülesandeid teostavad organisatsioonid võiksid samuti avalike ülesannete käigus teatavaks saanud informatsiooni avaandmetena avaldada. Arvatavasti on takistuseks teadmatus ning puudub tehniline tugi avaandmete avaldamise protsessi toetamiseks. Täiendavalt lisati, et asutustel on enda sisemised rakendused, infosüsteemid, mis ei ole riigi infosüsteemi osad, kasutuses. Ka nendes sisalduvad andmed võiksid olla läbipaistvuse huvides avalikustatud. Mitmed ministeeriumid maksavad organisatsioonidele erinevaks otstarbeks toetusi. Kui see informatsioon oleks avaandmetena ja ühtsel struktuuril avaldatuna kättesaadav, saab info kokku viia äriregistri andmetega ja vaadata, kas Eestis on selliseid organisatsioone, ettevõtteid, kelle tegevus on rahastatud peamiselt erinevatest toetustest. Lahenduse abiga on võimalik kontrollida ka toetuste maksmise süsteemi ärakasutamist, et ühel eesmärgil ei kasutataks mitut toetusmeedet ega esitataks korduvalt sama kulutõendit.

4.2.4 Tagasiside avaandmete portaalile

Intervjuude käigus ei küsitud otseselt tagasiside avaandmete portaali arendamise kohta, kuid siiski mitmelt intervjuult tuli mitmeid ettepanekuid selle täiustamiseks.

Mitmel intervjuul tuli teemaks avaandmete portaali kasutajamugavuse puudumine. Kui midagi on kohustuslik kasutada, siis peab see olema mugav ja üheselt arusaadavalt arendatud. Lisaks kasutajamugavusele võiks portaalis olla kättesaadavad ka erinevate töötubade materjalid, juhised jms, mis puudutab avaandmete valdkonda Eestis. Samuti tehti ettepanek korraldada tagasiside küsimustik portaali külastajatele. Tagasiside raames võiks külastajatelt küsida, et mida portaalis tehti ja mis huvi pärast pöörduti. Koostada mõned lihtsad küsimused, mille vastamine ei võtaks kaua aega. See annaks lisaks kasutajate statistilisele numbrile ka infot, mis eesmärgil portaali kasutatakse ning mis on külastajate huvid. Info võiks olla kasulik erinevatele teabevaldajatele ning anda ülevaadet ka valdkondadest, mille osas avaandmete kasutamine on aktiivsem või vastupidi.

Mitte tehnilise taustaga inimeste vestlustest selgus, et nad küll vastutavad andmete avaandmete portaali viitamise eest, kuid neil on seda raske teha, sest puudub Githubi kasutamise oskus. Eelduslik on mõningane programmeerimisoskus ja varasem kokkupuude Githubiga, et saaks andmeid avaandmete portaali lisada.

4.2.5 Avaandmete tervikvaade

Analüüsi käigus selgus, et avaandmete portaalist ei saa kätte tervikpilti. Juhul, kui kellelgi on huvi avaandmeid kasutada, siis täpse info leidmine ja tervikpildi koostamine, et mida-millisest andmehulgast saab, on keeruline. Võiks olla võimalik filtreerida märksõna või metaandmestiku andmeväljana kõik andmehulgad, mis sisaldavad näiteks isikukoode. Kuigi isikukoode pole avaldatud, siis saaks teada kõik andmehulgad, mis seda sisaldavad. Avaandmete portaali juurde võiks olla loodud metaandmete visualiseerimiseks puu. Puu annab hea ülevaate, kui metaandmete hulk on mingi tasemeni piiratud, sest olukorras, kus juurelemendil on mitukümmend alamelementi, on pilti keeruline hoomata. Puu kasutamise pluss on see, et ta on inimloetaval kujul ning kasutaja saab ülevaate andmeväljadest. Kuna paljud teabevaldajad avaldavad avaandmeid XML kujul, siis on oluline, et nende andmete struktuur on masinloetavalt kättesaadav XSD kaudu. See annab andmete kasutajatele ülevaate, mis

andmeväljad on andmekogus olemas. Metaandmed ja nende avaldamine on oluline teema. Mitmed inimesed tõid intervjuudel välja, et see vajab palju ressursi, aga kas panustatav aeg ja raha toovad eeldusliku kasu. Seega need inimesed, kes intervjuudel tõid välja, et pole kindlust avaandmete kasutamise osas, ei näe ka otsest vajadust täiendavaid ressursse panustada, et selgust luua. See omakorda tekitab olukorra, kus võibolla andmete kasutamiseks on soov, aga arusaamatult esitatud avaandmed pole kasutatavad.

4.2.6 Avaandmete litsents

Avaandmetega käib kaasas märksõna litsents. Andmekogule tuleb määrata litsents, et kasutajal oleks kindlus avaldatud andmeid kasutada ka ärilistel eesmärkidel. Suurema osa avaandmete puhul on avaandmete portaalis märgitud litsents. Samas on võimalik ka avaandmeteni jõuda läbi teabevaldaja kodulehe. Probleem tekib siis, kui neid andmeid soovitakse kasutada, aga märkimata on jäetud litsents. Litsentsi märkimata jätmine võib pärssida huvi ja ka julgust andmeid kasutada. Intervjuudest selgus, et litsentside sisust ei olda teadlikud ning seda teemat ei tähtsustata, kuid on valdkondi, milles on litsentside teema väga tähtis (autoriõigused).

Suuresti puudub selgus ja arusaam, miks avaandmeid on vaja avaldada. See on seadusest tulenev kohustus, mida on vaja täita, kuid sellele lisaks peaks olema teada valdkonna selged eesmärgid ja väärtus. Kuna seaduse täitmise järgi otsest kontrolli pole olnud, siis ei ole ka avaandmete teemaga tegeletud. Sellest aastast on hakatud rohkem avaandmete teemal sõna võtma ja jälgitakse avaandmete avaldamist, mis on põhjustanud elavnemist valdkonnas. Avaandmete valdkonna kohustuse täitmisega tegeletakse osaliselt seetõttu, et see on üheks kohustuslikuks eeltingimuseks struktuurivahenditest rahastatavate IT arendusprojektide koostamisel ning vahendite taotlemisel.

4.2.7 Kokkuvõtte intervjuudest

Kokkuvõtvalt võib intervjuude analüüsi tulemusena öelda, et valdkonnas soovitakse rohkem vajaduspõhist lähenemist. Kõik on nõus ilma piiranguteta avaandmeid avaldama, kui selleks on huvi. Seega on oluline leida lahendus, kuidas täita seadusest tulevat kohustust avaandmete avaldamise osas nii, et sellele kuluvad ressursid oleksid kasutatud efektiivselt, samas kõikidele huvitatutele oleks avaandmed kättesaadavad. Kindlasti tuleb tegeleda teavitamisega ja võimaluste tutvustamisega avaandmete valdkonnas. Praktilised näited saavutatud tulemustest innustavad valdkonda panustama ja innovatsiooni tekitama.

5 Avaandmete avaldamise mõjude hindamise töövahend

Magistritöö eesmärgist lähtuvalt koostas töö autor avaandmete avaldamise mõjude hindamise prototüüp lahenduse, mille aluseks on töö teoreetilises osas tutvustatud mõjuhinnangu tegemise kohustus ja rahvusvaheline praktika. Loodud töövahend on küsimuste kogum, mis peaks pakkuma tuge, vähendama koormust ja aitama avaliku sektori organisatsioonidel otsustada, mis andmeid avalikustada. Avaandmete avaldamise mõjude hindamise küsimustikku tutvustati avaliku sektori avaandmete eest vastutavatele ametnikele tagasiside saamise eesmärgil. Prototüüpi tutvustati intervjuude käigus. Tegemist on lahendusega, mida saab edaspidi huvi ja vajaduse korral täiendada ning praktikas kasutusele võtta. Mugavaks küsimustiku näitamiseks ja tagasiside saamiseks loodi veebileht²¹, kus on võimalik küsimustele vastata. Terviklik küsimuste nimekiri on esitatud magistritöö lisas II.

5.1 Ülevaade avaandmete avaldamise mõjude hindamise töövahendist

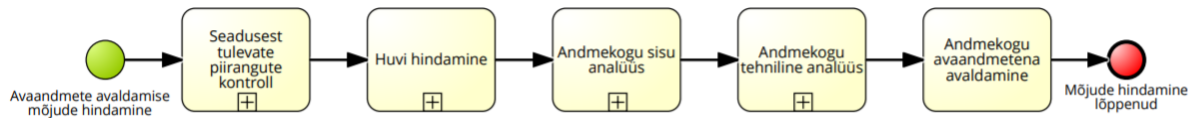
Avaandmete mõjuhinnangu koostamine on teabevaldajatele kohustus, kuid puudub teadmine, kuidas seda täna organisatsioonides tehakse ja kas tehakse. Ettevalmistatud küsimuste täitmise eesmärgiks on anda teabevaldajatele vastus, kas andmeid avaldada avaandmetena või mitte. Otseselt kontrolli täna ei ole, et organisatsioonid enda andmeid avaandmetena kättesaadavaks teeks. Andmeid, mida avaldada on kindlasti rohkem, kui need, mis jõuavad Euroopa Struktuurifondi vahenditest rahastatavate projektidena MKMi lauale, et siis avaandmete avaldamise tingimuste täitmist nõuda ja kontrollida. Selleks, et toetada teabevaldajaid avaandmete avaldamisel, on loodud avaandmete avaldamise mõjude hindamise töövahend. See aitab asustuste koormust vähendada ja lihtsustada otsustusprotsessi, kas andmeid võib avaldada või mitte. Küsimustele vastamine annab teabevaldajale peamised pidepunktid, millele tuleb mõelda enne andmete avaldamist avalikustamist. Kindlasti on palju rohkem detaile, millele peab enne andmete avaldamist mõtlema ja läbi analüüsima. Kõikides valitsemisalade valdkondades on spetsiifikat ja erisusi, mida üks töövahend ei suuda tervikuna katta. See tähendab, et loodud abistav küsimustik pidi olema niivõrd üldine, samas sisukas, et selle abil enamus organisatsioonidest saaks pidepunktid kätte, millele mõelda ning mis nüansse arvestada. Küsimustiku koostamisel on eeskujuks võetud Iirimaa avaandmete auditi tööriistas sisalduvad küsimused [15] ning avaandmete avaldamist reguleerivad nõuded AvTS-is. Küsimuste koostamisel on konsulteeritud avaandmete ekspertidega ning Andmekaitse Inspektsiooniga. Avaandmetega tegelevate spetsialistidega tehtud intervjuude käigus täiendavaid küsimuste lisamiseks ettepanekuid ei tehtud.

Prototüüpi, millele küsiti tagasiside, tutvustati 15 spetsialistile. Eesmärk ei olnud testida loodud prototüübi kasutajamugavust ning funktsionaalsust. Küsimustik sai loodud veebilehe vormis seetõttu, et seda oleks mugav näidata. Eesmärk oli saada küsimustele tagasiside, et seeläbi hinnata, kas loodud tööriistast oleks kasu ja võimaldaks selle abil teabevaldajatel selgust saada avaandmete avaldamise osas.

²¹ Avaandmete avaldamise mõjude hindamise töövahend: <http://kodu.ut.ee/~mariv/>

5.2 Ülevaade avaandmete avaldamise mõjude hindamise protsessist

Mõjude hindamise töövahendist tervikliku ülevaate andmiseks on autor valinud protsessi vaate kaudu lähenemise. Kõik töös sisalduvad protsessiskeemid on koostatud *Signavio Academic* protsesside koostamise töövahendiga.



Joonis 1. Avaandmete avaldamise mõju hindamise protsessi skeem

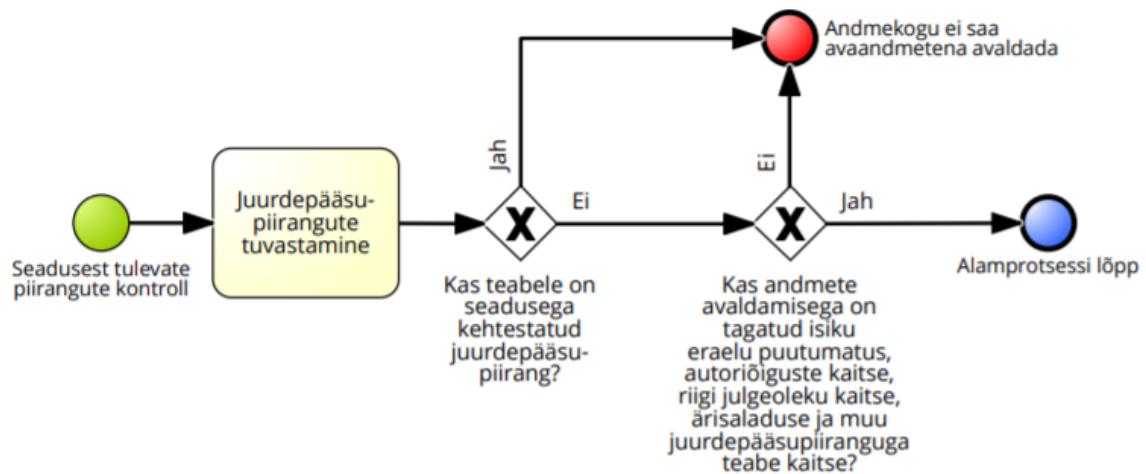
Pilt Joonis 1 annab ülevaate peamistest avaandmete avaldamise mõjude hindamise töövahendi osadest. Kõik osad (välja arvatud Andmekogu avaandmetena avaldamine) koosnevad eraldi alamprotsessidest. Loodud avaandmete avaldamise mõjude hindamise vahendist antakse ülevaade protsessi osade tutvustamise kaudu. Terviklik protsessi skeem on kättesaadav töö lisast III. Protsess on jaotatud viieks osaks:

- 1) Seadusest tulevate piirangute kontroll;
- 2) Huvi hindamine;
- 3) Andmekogu sisu analüüs;
- 4) Andmekogu tehniline analüüs;
- 5) Andmekogu avaandmetena avaldamine.

Järgnevalt annab autor ülevaate kõikidest protsessi osadest ning seal sisalduvatest küsimustest. Kõik küsimused ei sisaldu protsessi skeemis, sest valikvastustega lausega küsimused ei anna lisandväärtust protsessi, vaid on mõjude hindamisel teabevaldajale suunavaks abiks. Osad küsimused on olulised edaspidi avaandmete kasutajate jaoks ja annavad andmehulga kohta informatsiooni.

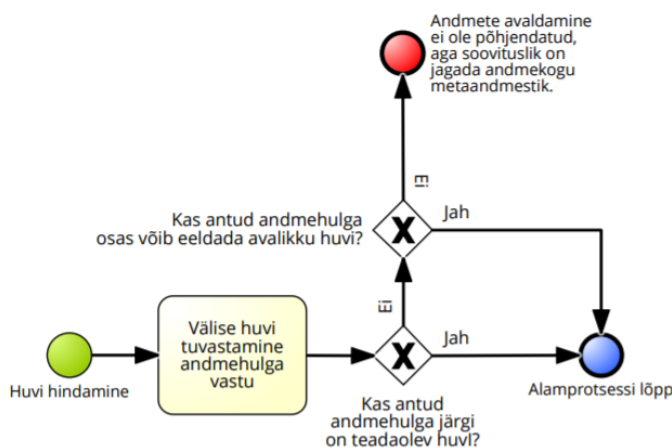
Lugejale infoks, et kõikide alamprotsessiskeemides tähistab sinine ring alamprotsessi lõppu, millele järgneb uus alamprotsess.

Esimene alamprotsess on seadusest tulevate piirangute kontroll ning kirjeldatud pildil Joonis 2. Protsessis on kaks küsimust, mis mõlemad proovivad välja selgitada, kas andmetele on seaduseandja piirangud lisanud. Esimene küsimus puudutab näiteks andmeid, mis on mõeldud vaid asutusesiseseks kasutamiseks ning lähtuvalt andmekogu põhimäärusest võib sellele olla määratud avaldamise piirang. Tasub tähele panna, et andmekogu põhimäärus võib seada ka teatud osale andmetest juurdepääsupiirangu. Sellisel juhul tuleb valida küsimusele vastuseks „Ei“ ning liikuda edasi järgmise küsimuse juurde. Järgmine küsimus on „Kas andmete avaldamisega on tagatud isiku eraelu puutumatus, autoriõiguste kaitse, riigi julgeoleku kaitse, ärisaladuse ja muu juurdepääsupiiranguga teabe kaitse?“. Kui vastaja valib vastusevariandiks „Ei“, siis ei ole võimalik andmeid avaldada avaandmetena. Oluline on tagada erinevate küsimustes nimetatud osapoolte kaitse. Juhul, kui andehulk piiranguid ei sisalda, on võimalik protsessis edasi liikuda ning jätkata teiste küsimuste vastamist.



Joonis 2. Seadusest tulevate piirangute kontroll

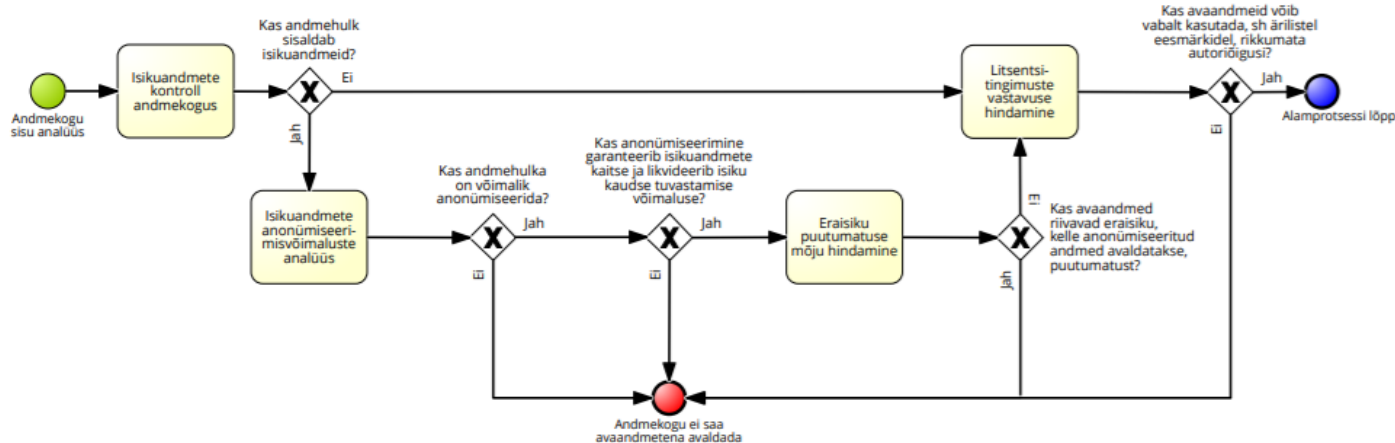
Kui seadusega pole andmetele seatud avaldamisena avaldamise osas piiranguid, liigub protsess edasi huvi hindamise osasse, mis on kajastatud pildil Joonis 3. Huvi hindamise osas on vaja teabevaldajal analüüsida, kas avaldamise kasutamise osas on väline huvi olemas ning kas seda huvi võib eeldada. Väline huvi all on mõeldud erinevaid osapooli, keda saab töövahendis täpsemalt ka määratleda. Sihtrühma osas saab teha valiku kodanike, akadeemia, ettevõtjate, avaliku sektori organisatsiooni või muu osapooli vahel (võimalik on märkida ka mitu osapooli). Autor on teinud valiku, et kui teabevaldaja vastab, et huvi andmete osas hetkel pole ning huvi pole eelduslikult ka tekkimas, siis ei ole otstarbekas avaldamisena andmeid avaldada, kui see eeldab täiendava rahalise ressursi panustamist vana infosüsteemi uuendamisel. Sellisel juhul peaks tulema küsimuste vastajale soovitus, et avaldamise metaandmed on vaja avalikustada avaldamise portaalil. Kui edaspidi peaks tekkima huvi andmeid masintöödeldavalt kasutada, siis ta saab metaandmete abil info, mis andmeid reaalselt peaks võimalik olema kasutada. Seejärel saab huviline teha andmete osas päringu, et need avaldataks. Teabevaldajale on etteantud aeg, mille jooksul peavad avaldamisena andmed kättesaadavaks tehtama. See lähenemine annaks valdkonda vajaduspõhisust, mida mitmed spetsialistid intervjuude käigus pidasid vajalikuks rõhutada. Autor on teadlik, et see lähenemine on vastuolus AvTS-iga.



Joonis 3. Huvi hindamine

Huvi hindamise järgselt tuleb teabevaldajal täita küsimused, mis on seotud andmekogu sisu analüüsiga ning keskenduvad isikuandmetele ja eraelu puutumatusele. Küsimused on näha pildil Joonis 4. Esimese sammuna tuleb teha isikuandmete kontroll andmekogus ja mõjude hindamise läbiviijalt küsitakse, kas andmekogu sisaldab isikuandmeid. Kui andmehulk sisaldab

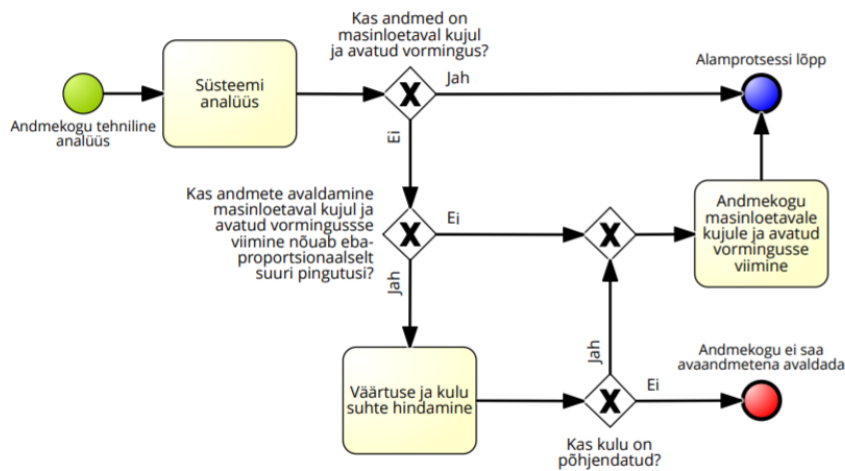
isikuandmeid, küsitakse vastajalt, kas neid on võimalik anonümiseerida. Pseudonümiseerimine on protsessist ja mõjude hindamise küsimustikust välja jäetud, sest pseudonümiseeritud andmed on võimalik tagasi isikustada ning täielikult pole tagatud eraelukaitses. Seejärel tuleb teabevaldajal analüüsida, kas anonümiseeritud andmete avaldamine garanteerib isikuandmete kaitses ning puudub kaudne isiku tuvastamise võimalus. Kui seda pole võimalik tagada, siis ei ole lubatud andmeid avaandmetena avaldada. Kui see on võimalik tagada, tuleb hinnata andmete avaldamise mõju. Protsessi „Seadusest tulevate piirangute kontroll“ osas küsitakse teabevaldajalt üldiselt eraelu puutumatuse kohta ning kui seal tuleb vastuseks, et avaandmetena isikuandmete avaldamine rikub eraelu puutumatust, siis ei tohi andmeid avaandmetena avaldada. Käesoleva etapi küsimus on seotud konkreetsemalt mõju hindamisega eraelu puutumatuse osas. See tähendab, et kui avaandmetena avaldatakse andmeid, mis on anonümiseeritud, siis tuleb hinnata, kas andmete avaldamisest võib tekkida isikutele mingisugune mõju. See mõju ei pruugi olla negatiivne, vaid võib olla ka positiivne. Juhul, kui vastaja on hinnanud, et eraisiku puutumus on tagatud, jõuab ta litsentsi tingimuste juurde. Vastajalt küsitakse, kas avaandmeid võib vabalt kasutada, sh ärilistel eesmärkidel, rikkumata autoriõigusi? Avaandmete puhul on oluline, et neile saaks määrata litsentsi, mis võimaldab andmeid vabalt taaskasutada, töödelda ja levitada. Kui see pole võimalik, siis ei ole kasu avaandmetena andmete avaldamisest. Soovituslik litsents avaandmete portaali eestvedajate poolt on *Creative Commons* 4.0 litsents, kuid see pole kohustuslik ning teabevaldaja võib määrata ka teise sobiva litsentsi. Hetkel on enim kasutatud *Creative Commons* 3.0 litsents.



Joonis 4. Andmekogu sisu analüüs

Kui sisuline osa on suuresti hinnatud ja on selge, et andmeid on võimalik avaandmetena avaldada, nad ei sisalda isikuandmeid, nende vastu on huvi ning on võimalik määrata litsents, mis võimaldab avaandmete taaskasutamist erinevatel eesmärkidel, jõuab protsess tehnilise poole juurde. AvTS ütleb, et avaandmeid peab masinloetavalt avaldama, kui see on võimalik ega nõua teabevaldajalt ebaproportsionaalselt suuri pingutusi. Ebaproportsionaalselt suure pingutuse hindamine on subjektiivne, ometi tuleb seda teha. Protsessis on teabevaldaja jõudnud viimasesse etappi, milleks on andmekogu tehniline analüüs. Pildil Joonis 5 on näha, et esmalt on vaja teada, kas andmed on juba masinloetaval kujul ja avatud vormingus. See tähendab, et infosüsteem on arendatud selliselt, et avaandmete avaldamise kohustusega on infosüsteemi arendusprotsessi algusest peale arvestatud ning võimekus avaandmeid avaldada on loodud. Kui seda võimekust ei ole ning masinloetavalt avaandmete avaldamine eeldab tööd, tuleb hinnata, kas see nõuab ebaproportsionaalselt suurt pingutust. Selle analüüsimisel tuleb kaaluda erinevaid aspekte, sh ka varasemalt läbikäidud avaandmete eeldatava huvi olemasolu osa. Kui küsimustikule vastab äripoole inimene, siis tal ei pruugi olla teadmist, mis kulud võivad kaasneda avaandmetena andmehulga avaldamise korral. Seetõttu on oluline konsulteerida IT-juhi või muu IT spetsialistiga, kes aitab hinnata tehtava kulu ja väärtuse suhet. Kui kulu ei ole

põhjendatud, siis avaandmetena andmeid ei pea avaldama. Üldiselt ei tohiks see avaandmete avaldamise juures takistuseks tulla. Intervjuude käigus suurem osa töid välja, et tegemist pole keeruka ega kalli protsessiga. Pigem jääb alati avaandmete avaldamine muude põhjuste taha, milleks võib olla äripoolte teadmatus või soovimatus valdkonnaga tegeleda.



Joonis 5. Andmekogu tehniline analüüs

Pildil Joonis 1 on näha, et pärast tehnilist analüüsi pole protsess lõppenud. Viimane viies osa on üldine ning sisaldab info sisestamist. Küsimustik on autori poolt loodud eesmärgiga, et kui seda hakatakse tulevikus rakendada, siis see võiks olla osa avaandmete portaalist. Portaali sisse logides saab töövahendi abil infot salvestada ning lõpus olevad üldised andmed lisatakse automaatselt portaali metaandmete väljale, et neid edaspidi dubleerivalt sisestama ei peaks. Üldine info puudutab andmete valdkonna sisestamist. See on oluline andmete filtreerimiseks ja otsimiseks. Lisaks on seal võimalus sisestada märksõnad, mille abil on võimalik vajalikku infot portaalist leida. Üldise info alla kuuluvad veel teabevaldaja ehk asutuse nimi, andmehulga koondnimi ehk pealkiri. Andmetest ülevaate saamiseks on võimalik lisada lühikokkuvõtte andmekogust. Juhul kui küsimustele vastaja saab varasemalt info, et andmehulk ei kuulu avaandmetena avaldamiseks, siis ei ole vaja üldist infot sisestada. Üldise info sisestamine on vajalik, kui on soov andmeid avaandmetena avaldada ja piirangud selle teostamiseks puuduvad.

5.3 Tagasiside avaandmete avaldamise mõjude hindamise töövahendile

Intervjuudel küsiti spetsialistidelt, kas nad näevad antud tööriistas väärtust nende igapäevases töös. Tagasiside oli erinev. Üldiselt nähti loodud avaandmete avaldamise mõjude hindamise töövahendis väärtust. Töövahendit nähti eelkõige kasulikuna andmekogude peakasutajale ehk neile äripoole esindajatele, kelle andmeid ei ole avaandmetena avaldatud ja kes pole avaandmete valdkonnaga varasemalt kokku puutunud. Küsimustele vastamine annab esmase sisendi otsuse tegemisel, kas andmete avaldamine on võimalik. Rohkem väärtust loob vahend siis, kui küsimustele vastuste valimine annaks täitjale kohehelt selgitavat infot lähtuvalt vastusest või annaks konkreetse soovitus, mida edasi teha või millele mõelda. Ehk kui vastatakse küsimustele „Kas andmehulka on võimalik anonümiseerida?“, siis antakse soovitus ja juhised, kuidas anonümiseerimine käib ja mis selleks vaja teha on. Käesoleva magistr töö skoopt soovitusete koostamine ei mahtunud. Autori hinnangul on soovitusete vajalikud, kui antud töövahend võetakse realselt kasutusse. Siis on võimalik erinevate valdkonna spetsialistidega koostöös parimad ja vajalikumad soovitusete kokku koguda ja kaardistada.

Arvati, et kui küsimustik ei anna konkreetset ja selget vastust ning põhjendust, kas andmeid võib avaldada või mitte, siis ei ole sellest kasu. Igapäevases on andmetega tegelevatel spetsialistidel soov saada konkreetset vastust küsimusele „Kas neid andmeid võib avalikustada?“. Intervjuudel toodi välja, et kui seda küsida, siis kunagi ei anta konkreetset vastust ning vastus langeb alati konkreetsele spetsialistile. Vastust aga ei soovita selles kontekstis võtta. Loodud töövahend ei ole ka mõeldud konkreetse vastuse ütleamiseks, vaid erinevate oluliste nüansside läbimõtleamiseks, et jõuda otsuseni. Tänane AvTS § 29 lõikes 3 nimetatud loetelu ei toeta avaandmete avaldamise eest vastutavaid spetsialiste. Juhul kui riik soovib sarnaselt Iirimaale avaandmete mõjude hindamise töövahendit kasutusele võtta, siis peab see olema erinevate osapoolte koostöös välja töötatud selliselt, et spetsialist saaks töövahendist saadud vastusele viidata, kui keegi tuleb kontrollima ja esitama küsimusi, miks andmeid on avaldatud.

Osaliselt öeldi, et ei taheta täiendavaid küsimustikke täita ning seda nähakse tüütu kohustusena. Rohkem aga öeldi intervjuudel, et tüütu kohustusena näevad avaandmete valdkonda need, kellel on õigel ajal infosüsteemi arenduse käigus avaandmete avaldamisele jäänud mõtlemata ning nüüd on vaja tagantjärele teemaga tegeleda. Kuigi tagasiside oli suuremas osas positiivne, siis väikese intervjuude mahu pealt ei ole võimalik üldistust teha.

Hetkel on kõik mõjude hindamise töövahendi küsimused kuvatud täitjale, kodulehe vaatajatele. Prototüüp lahendusena on see loodud selliselt seetõttu, et saaks intervjuude käigus võimalikult mugavalt tervikpilti näidata ja küsimustele tagasiside küsida. Autori eesmärk mõjude hindamise töövahendi praktikas rakendamisel oleks see, et täitjale avanevad küsimused ühekaupa. Lähtuvat vastatud küsimustest avaneb järgmine küsimus või soovitus lähtuvalt eelmisest vastusest. Inimestel võib tekkida huvi, millest mõjude hindamise töövahend tervikuna koosneb. Selle jaoks peaks looma juhise koos soovitusetega avaandmete avaldamise osas ning lisama sinna juurde küsimustiku, mille abil on võimalik mõjude hindamist läbi viia.

Arvati, et loodud töövahend aitaks kaasa valdkonna tervikpildi saamiseks. Võiks arvata, et enamus teavad täpselt, mis andmeid, missuguse andmekoosseisuga nende valitsemisalades on, kuid tegelikkuses see riigiülesest erineb. Intervjuudest selgus, et osaliselt on infot kaardistatud, kuid kindlasti oleks vaja infot põhjalikumalt. Sooviti, et saaks lingi küsimustega edastada allasutustele, kus peakasutajad küsimustele vastaks ning ministeeriumi IT osakond saaks endale vastustena sisendi.

6 Tulemused

Viimase mõne kuu jooksul on avaandmete avaldamine hoogustunud. Lõputöö ettevalmistuste tegemisel ja teema fookuseerimise hetkel, oli Eesti avaandmete portaalis ca 80²² andmehulka. Lõputöö esitamise ajaks²³ on andmekogude arv kahekordistunud, jõudes 172 andmehulgani. See tähendab, et asutustel on olemas võimekus avaandmeid avaldada. Selleks, et ka edaspidi andmete avaldamise hoog säiliks, on vaja protsessi lihtsustada avaliku sektori organisatsioonide jaoks. Magistritöö raames selgus, et üheks võimalikuks lahenduseks on avaandmete avaldamise mõjude hindamise töövahendi kasutusele võtmine. See aitab andmekogude peakasutajatel läbi mõelda, kas tema vastutada olevaid andmeid võib ja saab avaandmetena avaldada. Autor on nõus, et käesolevat vahendit tuleks täiendada lähtuvalt uurimuse käigus saadud tagasisidest.

Töö tulemusena selgus, et avaandmete valdkonna arenguks on vaja osapoolte teadlikkuse tõstmist. Kuigi tegemist on seadusest tuleva kohustusega, siis ei olda väga alati seda täitma. Jagama peab kogemusi, kus on avaandmete kasutamine aidanud mõnda probleemi lahendada või lihtsustada ja/või avaandmed on osaks mõnes rakenduses. Märksõnale „avaandmed“ väärtuse andmine peab toimuma läbi reaalsete näidete jagamise. Valdkonna arengusse saaksid panustada andmete kasutajad, näiteks ülikoolid, ettevõtjad, kodanikud. Kui oleks teada nõudlust andmete järgi, siis ollakse avatumad avaandmetega tegelemisel ja nende avaldamisel. Soovitakse näha, et avaandmeid kasutataks, nii oleks andmete avaldamise töö väärtus.

Oluline on tegeleda avaandmete kasutatavuse jälgimisega. Uurima peaks, kui palju kasutatakse avaandmeid masinloetaval kujul. Kui avaandmed on kättesaadavaks tehtud API-de kaudu, siis saab sinna lisada automaatse jälgimise. Täna ei ole teada, kui paljud on selle lisanud ja avaandmete kasutamist jälgivad.

Valdkonnas ühtsuse saavutamiseks oleks vaja standardit, mis annab konkreetseid suunised teabevaldajatele. Uute arendusprojektide kavandamisel oleksid selged juhised avaandmete avaldamisest abiks, et nendega saaks nõuete kirjeldamisel kohe arvestada. Täna asutustele kasutada olevad materjalid on osaliselt aegunud ning vajaksid uuendamist ja ajakohastamist. Dokumentide uuendamisel tasub mõelda ühe tervikmaterjali loomise peale, millest saaks avaandmetega tegelevad inimesed informatsiooni kätte.

Laiemalt ei osata näha väärtust avaandmetest, sest ei mõisteta ühe andmekogu avaldamise olulisusest ja väärtusest. Täna vaatavad kõik asutused enda andmeid ning ei mõtle sellele, et kui avaandmed tehakse kättesaadavaks näiteks teise struktuuriga, siis saaks teise valitsemisala andmekoguga kokkupanekul luua rakenduse, mis aitaks säästa aega. Laiem vaade valdkonnale on puudu ega nähta võimalust, et erinevate avaandmete koostoimes võiks nende pealt luua väärtuspakkuva toote või teenuse. See eeldab teadlikkuse tõstmist ning regulaarset avaandmetega tegelemist.

Avaandmete avaldamine pole kindlasti ainult avaliku sektori võimalus. Panustada saavad ka kodanikud, eraettevõtted enda avaandmete jagamisega. Näiteks on NASA-l Eesti kohta õhukvaliteedi ja lumevaiba avaandmed olemas, mida võibolla soovitaks kasutada, aga kui nende olemasolust ei teata, siis ei osata neid ka võibolla otsida. Samuti oleks vaja analüüsida seda, mis roll avaandmete avaldamisel on asutustel, kes täidavad avalikke ülesandeid, aga tegemist ei pruugi alati olla avaliku sektori organisatsioonidega.

Avaandmete laialdane ja masinloetavalt mugavalt kättesaadavaks tegemine edendab rakenduste loomist. Avaandmete portaali kaudu on viidatud, et rakendusi on loodud 18²⁴. Mitmed

²² 2019 aasta alguse seis.

²³ 16. mai 2019 aasta seisuga on 172 andmehulka avaandmete portaalis, <https://opendata.riik.ee/andmehulgad/>

²⁴ 16. mai 2019 seisuga avaandmete portaalis, <https://opendata.riik.ee/rakendused/>

intervjueeritavad arvavad, et neid rakendusi on oluliselt rohkem, kuid neid ei ole ära markeeritud avaandmete portaalis.

Magistritöö raames loodud avaandmete avaldamise mõju hindamise töövahend, millele küsiti valdkonna spetsialistidelt tagasiside. Tagasiside loodud töövahendile oli enamasti positiivne ja soovitakse lahendust kasutusele võtta. Samuti tehti ettepanekuid selle täiendamiseks ja arendamiseks.

Avaandmete valdkonna jätkuvaks arenguks pakub autor välja järgmised tegevused:

- Tehnoloogilises raamistikus kokku leppimine, kuidas ühtselt andmeid kasutada, et neid oleks võimalik soovijatel mugavalt taaskasutada. See eeldab valdkonda reguleerivate juhiste kirjeldamist ja uuendamist. Kaasajastatud juhised ja ühtsed reeglid võimaldavad andmekogu omanikel kindlustunnet ja teadmist, kuidas õigesti andmeid avaldada.
- Oluline on kokku leppida ühtses andmevahetuses infosüsteemide vahel ja ühtlustatud andmekirjelduses. Ühtne metaandmete kirjeldamise standard, mida kõik jälgivad ja kasutavad, aitab muuta avaldatud avaandmeid selgemini mõistetavaks.
- Avaandmete avaldajad võiksid ise valdkonda panustada ja välja pakkuda võimalusi, kus ja mille jaoks nende avaandmeid kasutada. See võiks suurendada avaandmete kasutamist, mis omakorda edendaks avaliku huvi korral teisi teabevaldajaid andmeid avaldama ja kättesaadavaks tegema.

Töö teoreetilises osas kirjeldatud probleemid, mis on seotud avaandmete avaldamisega, on osaliselt aktuaalsed ka Eesti avalikus sektoris. Eelnevalt teoorias nimetatust toodi ka Eestis välja seda, et avaandmete valdkond ei ole prioriteetne ning seetõttu ei nähta vajadust aega ja raha panustada. Olenemata kohustusest ei soovita valdkonda panustada. Sarnaselt varasemate uuringute käigus selgunud infoga, kardetakse ka Eestis eksida isikuandmete kaitse reeglite vastu ning vajatakse tuge otsustamiseks, kas ja mis isikuandmeid võib avaandmetena avaldada. Kui mujal riikides kardetakse, et avaandmete avaldamisega kaasnevad turvariskid, siis seda Eestis ei nimetatud. Pigem ei seostatud teemat avaandmetega. Eesti kontekstis ei selgunud uurimuse käigus, et andmete avaandmetena avaldamist takistaks kehv andmekvaliteet.

7 Kokkuvõte

Magistritöö eesmärk oli luua töövahend avaandmete avaldamisele eelneva mõjuhinnangu tegemiseks. Avaandmete avaldamise eel on kohustuslik läbi viia avaandmete mõjuhinnang²⁵. Mõjuhinnang on vaja teha andmekogude põhimääruse muutmisel ja uue andmekogu loomisel, kui on soov andmeid avaldada avaandmetena. See kohustus ei ole täna rakendunud ning vastavaid mõjuhinnanguid pole suuresti teostatud. Loodud praktiline vahend aitab teabevaldajat mõjude hindamise protsessis ja järgneva otsuse tegemisel, kas andmekogus sisalduvaid andmeid võib avaldada või mitte ning mis on vajalikud eelnevad tegevused andmete avaldamiseni jõudmisel.

Mõistmaks, miks Eesti, kui muidu edukas digiriik, avaandmete vallas kehval tasemel on, viis autor läbi uurimuse arendusprojektide nõuete täitmise osas. Valdkonna paremaks mõistmiseks tegi autor täiendavad intervjuud avaandmetega tegelevate avaliku sektori spetsialistidega. Valdkonna spetsialistid toetavad avaandmete avaldamist, kuid soovivad näha selle tegevuse mõju. Intervjuud andsid ülevaate, mis põhjustel avaandmeid pole avaldatud. Üheks suureks põhjuseks on teadlikkuse puudumine, sest kohustus avaandmeid masinloetavalt avaldada on Eestis juba alates 2015. aastast.

Magistritöö eesmärgi saavutamiseks sõnastati kolm uurimisküsimust, mis said töö käigus vastuse. Esimene uurimisküsimus keskendus avaandmekogude väiksele mahule. Olukord on magistritöö kirjutamise käigus järjest paranenud (projektide analüüsi tulemusena võetud teabevaldajatega ühendust). Kui töö kirjutamise alguses oli avaandmete portaali kaudu kättesaadav ca 80 andmekogu, siis töö esitamise ajaks on neid 172. Põhjuseks toodi välja ka teadmatust. Mitmed asutused olid avaldanud avaandmeid enda portaalide kaudu, kuid puudusid viited riiklikule avaandmete portaalile, sest ei teatud, et seda peab tegema.

Teise uurimisküsimusena oli eesmärk teada saada, kuidas võiks edendada avaandmete avaldamist. Analüüsimisel selgus, et tuntakse puudust konkreetsetest juhistest ja suunistest, mida järgida. Kuna täna on valdkonnas väga erinevat lähenemist, kuidas avaandmeid avaldada ja mitmendale tärnile avaldatud andmed peaksid vastama, siis kohustuslikud juhised toetaksid valdkonna arengut. Lisaks aitaks avaandmete edendamisele kaasa teadlikkuse tõstmine. Oluline on jagada kogemusi tehtud lahendustest, mis kasutavad avaandmeid. Paljud spetsialistid ei näe valdkonda panustamisest kasu, sest avaldatud avaandmeid pole teadaolevalt kasutatud. Mitmed intervjuueeritavad rõhutasid väliste osapoolte tähtsust, ka näiteks akadeemia võiks rohkem avaandmeid küsida, mis omakorda motiveerib avaandmete avaldamisega tegeletama. Väliste osapoolte huvi avaandmete vastu on ka võimalus juhtkonnale valdkonna olulisuse rõhutamisel.

Kolmanda uurimisküsimusega sooviti uurida, kas avaandmete mõjuhinnangu koostamise töövahend võiks toetada ja edendada avaandmete avaldamist. Magistritöö raames küsiti loodud töövahendile tagasiside ning analüüsiti, kas ja kuidas pakutud lahendust edaspidi praktikas rakendada. Avaandmete avaldamise mõjude hindamise töövahendit pidasid enamik intervjuueeritavatest vajalikuks ja kasulikuks abivahendiks andmekogude peakasutajatele.

Loodud töövahendile on võimalik laiemat tagasiside valitsemisalade esindajatelt küsida ning lähtuvalt sellest on võimalik seda edasi arendada ja kasutusele võtta. Avaandmete valdkonna poliitikakujundajale on töövahendit ja sellele saadud tagasiside tutvustatud ning plaanis on töövahendiga edasi tegeleda ja praktikas rakendada.

²⁵ Avaliku teabe seadus, <https://www.riigiteataja.ee/akt/115032019011?leiaKehtiv>

8 Viidatud kirjandus

- [1] Euroopa komisjon. 2018. aasta digitaalmajanduse ja -ühiskonna indeks (DESI), Eesti riigiaruanne. http://ec.europa.eu/information_society/newsroom/image/document/2018-20/ee-desi_2018-country-profile-lang_4AA7568B-C99C-9BA3-F9FA95B265D75C8B_52352.pdf (3.03.2019).
- [2] United Nations. E-Government Development Index – Estonia. *UN EGDI Survey*, 2018. <https://publicadministration.un.org/egovkb/en-us/Data/Country-Information/id/57-Estonia> (11.04.2019).
- [3] World Wide Web Foundation, 2017. OpenData Barometer Global Report Forth Edition, Open Data Barometer. <https://opendatabarometer.org/doc/4thEdition/ODB-4thEdition-GlobalReport.pdf> (27.04.2019).
- [4] Riigikogu. Avaliku teabe seadus, Riigi Teataja. <https://www.riigiteataja.ee/akt/115032019011?leiaKehtiv> (14.05.2019).
- [5] Majandus-ja Kommunikatsiooniministeeriumi riigi infosüsteemide osakond, 2014, Eesti avaliku teabe masinloetava avalikustamise roheline raamat. Tallinn: Vabariigi Valitsus.
- [6] Riigi Infosüsteemi Amet, PricewaterhouseCoopers Eesti, 2016, Andmekvaliteedi tagamise juhend andmekogu omanikele, juhend. https://www.ria.ee/sites/default/files/content-editors/publikatsioonid/andmekvaliteedi_tagamise_juhend_andmekogu_omanikele.pdf (12.05.2019).
- [7] Tauberer J. Open Government Data: The Book, Second Edition: 2014.
- [8] Janssen M., Charalabidis Y., Zuiderwijk A. Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management* 29:4, 2012, 258-68.
- [9] Houghton J. 2011. Costs and Benefits of Data Provision, Report to the Australian National Data Service. https://www.ands.org.au/__data/assets/pdf_file/0004/394285/houghton-cost-benefit-study.pdf (14.05.2019).
- [10] Pollock R., Rogers K. Open Data Handbook, How the UK Government Saved £4 million in 15 minutes with Open Data. <http://opendatahandbook.org/value-stories/en/saving-4-million-pounds-in-15-minutes/> (14.05.2019).
- [11] Laursen, K., Salter A. Open For Innovation: The Role of Openness In Explaining Innovation Performance Among U.K. Manufacturing Firms. *Strategic Management Journal* 27: 2006, 131 – 150.
- [12] Vabariigi Valitsus. Eesti Infoühiskonna Arengukava 2020. https://www.mkm.ee/sites/default/files/eesti_infouhiskonna_arengukava_2020_ja_kuberturvalisuse_strateegia_2019-2022.pdf (10.04.2019).
- [13] Magalhaesa G., Roseirab C. Open government data and the private sector: An empirical view on business models and value creation. *Government Information Quarterly*. 2017. <http://dx.doi.org/10.1016/j.giq.2017.08.004> (03.03.2019).
- [14] Conradie P., Choenni S. On the Barriers for Local Government Releasing Open Data, *Government Information Quarterly* 2014. <https://www.sciencedirect.com/science/article/pii/S0740624X14000513?via%3Dihub> (04.03.2019)

- [15] Barry E., Bannister F. Barriers to Open Data Release: A View from the Top. Information Polity. 2014 doi:10.3233/IP-140327. (04.03.2019).
- [16] Young M., Yan A. Civic Hackers' User Experiences and Expectations of Seattle's Open Municipal Data Program, *Proceedings of the 50th Hawaii International Conference on System Science (HICSS-50)*, Hawaii, USA. 2017.
- [17] Tim Berners-Lee, 2006. Linked Data. <https://www.w3.org/DesignIssues/LinkedData.html> (20.04.2019).
- [18] Government Reform Unit of the Department of Public Expenditure and Reform. Open Data Audit and Publication Planning. <https://data.gov.ie/sites/default/files/files/Briefing%20Note%20on%20data%20audits%20-%20Final.pdf> (5.05.2019).
- [19] Riigi Infosüsteemi Amet. RIHA koolitusmaterjalid 2019. <https://moodle.ria.ee/mod/page/view.php?id=648> (14.05.2019).
- [20] Andmekaitse Inspektsioon, 2019, Mõjuhinnaang on riskide kaardistamine ja hindamine. <https://www.aki.ee/et/uudised/uudiste-arhiiv/mojuhinnang-riskide-kaardistamine-ja-hindamine> (5.05.2019).
- [21] REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, Official Journal of the European Union. http://publications.europa.eu/resource/cellar/3e485e15-11bd-11e6-ba9a-01aa75ed71a1.0006.03/DOC_1 (5.05.2019).
- [22] European Digital Rights Association, Foundation for Information Policy Research. An Introduction to Data Protection. The European Digital Rights papers. https://edri.org/files/paper06_datap.pdf (19.04.2019).
- [23] Euroopa Andmeportaal, 2018. Protecting data and opening data. <https://www.europeandataportal.eu/et/highlights/protecting-data-and-opening-data> (12.05.2019).
- [24] World Wide Web Foundation, 2018, Open Data Barometer: Leaders Edition, From Promise to Progress. <https://opendatabarometer.org/doc/leadersEdition/ODB-leadersEdition-Report.pdf> (18.04.2019).
- [25] Government Reform Unit, Department of Public Expenditure and Reform, 2017, Open Data Strategy 2017 – 2022. [https://data.gov.ie/sites/default/files/files/Final%20Strategy%20online%20version\(1\).pdf](https://data.gov.ie/sites/default/files/files/Final%20Strategy%20online%20version(1).pdf) (5.05.2019).
- [26] Department of Public Expenditure and Reform (Ireland), 2015. Open Data Technical Framework. <https://data.gov.ie/pages/opendatatechnicalframework> (20.04.2019).
- [27] Government of Canada, 2017, About the Open by Default Pilot. <https://open.canada.ca/en/about-open-by-default> (12.05.2019).
- [28] Government of Canada, Do-it-Yourself Open Data Toolkit. <https://open.canada.ca/en/toolkit/diy> (12.05.2019).
- [29] Riigi Infosüsteemi Amet, Struktuuritoetused 2014-2020. <https://www.ria.ee/et/struktuuritoetused/2014-2020.html> (12.05.2019).

- [30] Riigi Infosüsteemi Amet. (2019). Jooksev voor: avaandmete ja andmeanalüüsi kasutuselevõtu edendamine. <https://www.ria.ee/et/struktuuritoetused/2014-2020/avalikud-teenused/voor-avaandmed.html> (3.03.2019).
- [31] Denzin, N., Lincoln, Y. The Landscape of Qualitative Research. Thousand Oaks: Sage Publications. 2008.
- [32] Creswell, J. Research design: Qualitative, quantitative and mixed method approaches. Thousand Oaks: Sage Publications. 2003.
- [33] Flick, U. *Introducing Research Methodology: a Beginner's Guide to Doing a Research Project*. London: SAGE Publications. 2011.
- [34] Kvalitatiivne sisuanalüüs (2015), Kalmus V., Masso A., Linno M. <http://samm.ut.ee/kvalitatiivne-sisuanalyys> (10.05.2019).
- [35] Saldana, J. *The coding manual for qualitative researchers*. Los Angeles: SAGE. 2009.

9 Lisad

9.1 Lisa I Intervjuu kava

Üldine osa

- Mis lähenemine on valdkonnas avaandmetega? Kas avaldate kohustuslikke andmeid, mida avaliku teabe seadus kohustab või laiemalt?
- Kuidas andmete avaldamisse juhtkond suhtub?
- Missugune on huvi teie valdkonna andmete vastu? Kas ja kes andmeid soovib?
- Kas teie andmete pealt on loodud mingeid rakendusi, teenuseid? Kui jah, siis kas andmete kasutajatega on ka omavahelist suhtlust?

Avaandmed

- Mitme andmekogu andmed on avaandmetena avaldatud? Kas need on avaldatud masinloetavas formaadis? Kui kõik pole avaldatud, siis mis põhjustel?
- 5 täрни vaatest lähtuvalt, mis tasemel ja formaadis avaandmeid avaldate? Miks?
- Kas andmed on allalaaditavad või on API? Miks?
- Mitmed neist on viidatud avaandmete portaalis?
- Mis takistab avaandmete avaldamist? Mis on keerukus andmete avaldamise taga?
- Mis soodustaks andmete avaldamist?
- Mida oleks teil vaja, et mugavalt andmeid avaldada ja seadusest tulenevaid kohustusi täita?

Protsess

- Missugune on protsess teil avaandmete avaldamisega?
- Kui tihti te uuendate andmeid ja kuidas see on korraldatud?
- Kuidas metaandmeid avaldate? Kas need on masinloetavalt olemas?
- Kas peaks olema ühine üleriigiline reeglistik metaandmete koostamise ja uuendamise osas?
- Missugused ettepanekud on teil tänase avaandmete protsessi muutmiseks?
- Mis oleks teie arvates abiks andmete avaldamisel?

Turvalisus

- Kuidas olete kindlad, et mingeid andmeid ei tohi avaldada või just tohib avaldada? Mis protsess teil eelneb andmete avaldamiseks märksõnale turvalisus mõeldes?
- Kuidas on teil korraldatud anonümiseerimine?

Uued projektid

- Kui plaanite uut arendust hakata valdkonnas tegema, siis mis roll on ettevalmistamise protsessis märksõnal avaandmed?
- Kuidas peaks olema uute arendusprojektide puhul avaandmete teema käsitletud?
- Mis etapis avaandmed täna arendustel teemaks tulevad?
- Kas nende olemasolevate infosüsteemide puhul, mille andmeid ei ole avaandmetena avaldatud, aga seadusest ei tulene avaldamise osas piiranguid, siis kas neid on teil täna võimalik avaldada? Miks ei ole avaldanud?

Kas teil on veel midagi, mida soovite avaandmete teemal öelda?

9.2 Lisa II Avaandmete avaldamise mõjude hindamise töövahend

Avaandmete avaldamise eel on kohustuslik läbi viia avaandmete mõjuhindang. Avaandmete avaldamise mõju hindamise töövahendi eesmärk on aidata andmekogu peakasutajal seda teha. Samuti on loodud töövahendi eesmärgiks aidata teha otsust, kas andmeid võib avaldamisena avaldada või mitte. Loodud küsimused aitavad mõelda teemadele, mida enne andmete avaldamist analüüsida ning kaaluda, missugune on avaldamisega kaasnev mõju.

Küsimused:

- Kas teabele on seadusega kehtestatud juurdepääsupiirang? Jah/Ei
- Kas andmete avaldamisega on tagatud isiku eraelu puutumatuse, autoriõiguste kaitse, riigi julgeoleku kaitse, ärisaladuse ja muu juurdepääsupiiranguga teabe kaitse? Jah/Ei
- Kas antud andmehulga järgi on teadaolev avalik huvi? Jah/Ei
- Kas antud andmehulga osas võib eeldada avalikku huvi? Jah/Ei
- Kas andmehulk sisaldab isikuandmeid? Jah/Ei
- Kas andmehulka on võimalik anonümiseerida? Jah/Ei
- Kas anonümiseerimine garanteerib isikuandmete kaitse ja likvideerib isiku kaudse tuvastamise võimaluse? Jah/Ei
- Kas avaandmed riivavad eraisiku, kelle anonümiseeritud andmed avalikustatakse, puutumatust? Jah/Ei
- Kas avaandmeid võib vabalt kasutada, sh ärilistel eesmärkidel, rikkumata autoriõigusi? Jah/Ei
- Kas andmed on masinloetaval kujul ja avatud vormingus? Jah, Ei?
- Kas andmete avaldamine masinloetaval kujul ja avatud vormingusse viimine nõuab ebaproportsionaalselt suuri pingutusi? Jah/Ei

Üldised küsimused:

Mis avaliku teabe seaduse § 29 lõikest 3 tuleneva avaandmete avaldamise punkti alla kuulub avaldatav andmehulk?

1. riigi ja kohaliku omavalitsuse üksuste üldistava majandusstatistika ja majandusprognoosid;
2. kuritegevuse ja väärtegevuste üldistav statistika;
3. riigi- ja kohaliku omavalitsuse asutuste ning nende struktuuriüksuste põhimäärused;
4. riigi- ja kohaliku omavalitsuse asutustele esitatavate avalduste ja muude dokumentide vormid ning nende täitmise juhendid;
5. andmed ohu kohta inimeste elule, tervisele ja varale;
6. riigi- ja kohaliku omavalitsuse asutuste töötulemuste ja ülesannete täitmise aruanded;
7. avalik-õiguslike juriidiliste isikute tegevusaruanded ja tulude-kulude aruanded;
8. riigiasutuste ja kohaliku omavalitsuse üksuste ning kohaliku omavalitsuse asutuste eelarvete eelnõud ja eelarved ning nende täitmise aruanded;
9. andmed riigieelarve tulude laekumise kohta;
10. andmed keskkonnaseisundi, keskkonnakahjustuste ja ohtlike keskkonnamõjude kohta;
11. ministeeriumides valminud seaduste ja Vabariigi Valitsuse määruste eelnõud koos seletuskirjadega nende kooskõlastamiseks saatmisel ja Vabariigi Valitsusele esitamisel;
12. ministri ja kohaliku omavalitsuse organite määruste eelnõud koos seletuskirjadega enne nende vastuvõtmiseks esitamist;

13. kontseptsioonide, arengukavade, programmide ja muude üldise tähtsusega projektide eelnõud enne pädevatele organitele heakskiitmiseks esitamist, samuti vastavad heakskiidetud või vastuvõetud dokumendid;
14. riigi- või kohaliku omavalitsuse asutuse tellitud uuringud ja analüüsid;
15. andmed riigi- ja kohaliku omavalitsuse asutuste täitmata ametikohtade kohta;
16. andmed riigi ja kohaliku omavalitsuse üksuste poolt teostatavate ja teostatud riigihangete kohta;
17. andmed selle vara ja eelarveraha kasutamise kohta, mille riik või kohaliku omavalitsuse üksus on andnud üle riigi või kohaliku omavalitsuse üksuse asutatud või nende osalusega eraõiguslikele juriidilistele isikutele;
18. avalike ürituste kavad;
19. muudatused riigi- ja kohaliku omavalitsuse asutuste töös ja ülesannetes, mis on seotud isikute teenindamisega, vähemalt kümme päeva enne muudatuste rakendamist;
20. andmed riigi- ja kohaliku omavalitsuse asutuste juhtide vastuvõtuaegade kohta;
21. andmed kaubaturul valitsevas seisundis olevate, eri- või ainuõigust või loomulikku monopoli omavate äriühingute hinnakujunduse kohta;
22. andmed üldkasutatavate teenuste osutamise kohta, samuti muudatuste kohta teenuste osutamise tingimustes ja hindades enne selliste muudatuste rakendamist;
23. isikuandmete töötlemise eesmärk, ulatus ja viis, kolmandatele isikutele, sealhulgas teisele asutusele, isikuandmete edastamise ja avalikkusele kättesaadavaks tegemise ning isiku poolt enda andmetega tutvumise õigus ja kord;
24. teave avaandmete kättesaadavuse ja vajaduse korral litsentside kohta;
25. avaldatavad avaandmed ei kuulu eelpool nimetatud kohustuslike avaandmete hulka.

Kui vastaja valib viimase, siis lisaküsimus: Palun lisage täpsustus, mis avalike ülesannetega seotud avaandmetega on tegemist?

- Kes on andmekogu potentsiaalsed kasutajad? Sihtgrupi valik: kodanikud, tudengid, teadlased, ettevõtjad, teine avaliku sektori asutus, muu (avatud väli ja palve täpsustada)
- Mis ajast alates on andmekogus andmed? YYYY/MM/DD
- Mis ajaperioodi kohta käivaid andmeid plaanite avaldada? Vastusevariandid: Alates YYYY/MM/DD kuni YYYY/MM/DD;
- Kui tihti toimub andmete uuendamine? Vastusevariandid: Andmeid ei uuendata, jooksvalt, igas minutis, igas tunnis, igapäevaselt, kord nädalas, kord kuus, kord kvartalis, kord aastas
- Millal toimus viimane andmete uuendamine? YYYY/MM/DD
- Mis tärnile vastavas formaadis avaandmed avaldatakse? Variandid: 1-5 täрни

★

★★

★★★

★★★★

★★★★★

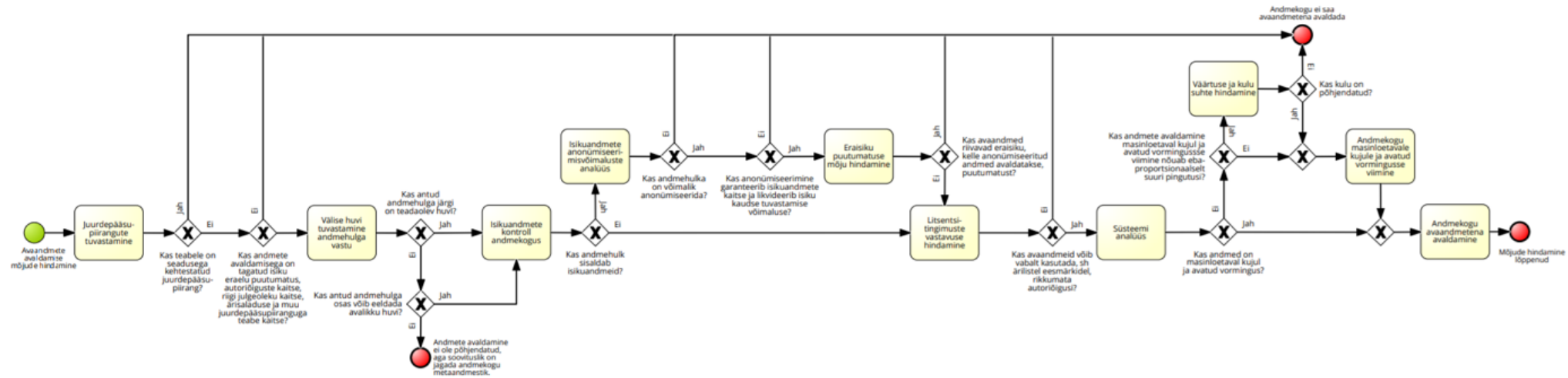
Lisainformatsioon:

- Teabevaldaja nimi
- Andmekogu pealkiri

- Andmekogu lühikirjeldus
- Mis valdkonna andmetega on tegu:
 - Avalik sektor, valitsus
 - Avalik turvalisus
 - Elanikkond ja ühiskond
 - Ehitus
 - Ettevõtlus
 - Haridus
 - Ilm
 - Kultuur
 - Kunst
 - Keskkond
 - Majandus ja rahandus
 - Piirkonnad, linnad
 - Põllumajandus, kalandus, metsandus
 - Rahandus, finants
 - Rahvusvahelised suhted
 - Riigiteenused
 - Sport
 - Teadus ja tehnoloogia
 - Tervis
 - Transport
 - Õigusemõistmine, õigussüsteem
 - Muu – palun täpsustada

Palun lisage märksõnad, mis kirjeldaks andmehulka: Vaba teksti väli

9.3 Lisa III Avaandmete avaldamise mõjude hindamise protsessi terviklik skeem



10 Litsents

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Mari Väljaots,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose „Avaandmete avaldamise mõjude hindamise töövahend“, mille juhendajad on Vambola Leping ja Ott Velsberg, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.
2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Mari Väljaots

16.05.2019